Kai Sauerwald, Matthias Thimm

Wissensrepräsentation und -verarbeitung

1 Einleitung

Wissensrepräsentation und -verarbeitung (engl. *knowledge representation and reasoning*) ist eine Teildisziplin der Künstlichen Intelligenz, die sich mit der formalen und expliziten Repräsentation von Information und Wissen, sowie der automatischen Anwendung von Schlussfolgerungsmechanismen zur Ableitung implizitem Wissens, beschäftigt. Besonders relevant bei dieser Definition und das Hauptunterscheidungsmerkmal zu vielen anderen Teildisziplinen der Künstlichen Intelligenz, insbesondere des *Deep Learnings*, ist hier, dass Wissen explizit durch formale Methoden dargestellt wird.

Die Repräsentation und Bereitstellung von Wissen ist eine für die menschlische Zivilisation ungemein wichtiger Aspekt. Das Niederschreiben und die Weitergabe wissenschaftlicher Erkenntnisse ist beispielweise zentral für den wissenschaftlichen Fortschritt und die Entwicklung jeglicher Technologie. In gleicher Weise gilt dies auch für alle anderen Aspekte der menschlichen Kultur wie Literatur und Poesie. Die Entwickung von Schrift, Buchdruck und zuletzt der digitalen Speicherung von Wissen sind für unsere heutige *Wissensgesellschaft* von höchster Wichtigkeit. Die Fähigkeit des Menschen in Schriftform vorliegendes Wissen zu verarbeiten, neue Schlussfolgerungen zu betreiben und diese wieder in Schriftform zu repräsentieren, erlaubt die verteilte, parallele und aufeinander aufbauende Bearbeitung unterschiedlichster Themen. Heutzutage steht mit dem *World Wide Web* (*WWW*) ein äußerst mächtiges Werkzeug zur (menschlichen) Wissensverarbeitung zur Verfügung und Webangebote wie Wikipedia bieten einen breiten Zugriff auf das Wissen der Welt.

Informationssysteme wie das WWW werden aber nicht nur von menschlichen Agenten benutzt. Immer häufiger finden sich dort auch künstliche Agenten, die im Auftrag ihrer menschlichen Klienten Aufgaben erledigen. Assistenzsysteme finden Antworten auf Fragen oder bestellen Produkte bei Online-Marktplattformen. Verschiedene Onlineangebote zur Planung von Reisen suchen und kombinieren Flugverbindungen verschiedenster Anbieter, um die beste Lösung für Anfragen zu ermitteln. Damit solche Systeme das vorliegende Wissen sinngemäß verarbeiten können, muss es in einer maschinenlesbaren Form vorliegen und bereitgestellt werden. Während der Aspekt der (technischen) Bereitstellung eher in die For-

Kai Sauerwald, Matthias Thimm, FernUniversität in Hagen
https://doi.org/10.1515/9783110218091-001

schungsbereiche der Informationssysteme und Datenbanken fällt, werden wir uns in diesem Kapitel eher mit grundlegenden Fragestellungen der Repräsentation und Nutzung von Wissen beschäftigen.

2 Grundlegende Konzepte

Das zentrale Werkzeug für das Forschungsgebiet der *Wissensrepräsentation und -verarbeitung* ist die *formale Logik*. Mit dem Begriff *Logik* wird allerdings nicht nur ein konkreter Formalismus bezeichnet, sondern eine ganze Familie unterschiedlichster Ansätze zur Repräsentation von Wissen und zur Formalisierung von Schlussfolgerungen. Im Allgemeinen besteht ein logisches System aus der Spezifikation einer *Syntax* und einer *Semantik*. Die Syntax definiert dabei die *Objektebene*, d. h., die Strukturen und Symbole, die in der entsprechenden Sprache wohlgeformte Ausdrücke darstellen. Die Semantik hingegen definiert die *Bedeutung* wohlgeformter Ausdrücke auf der *Metaebene*. Mithilfe syntaktischer Objekte wird das *explizite* Wissen in einer *Wissensbasis* repräsentiert, also solches Wissen (beispielsweise Fakten- oder Regelwissen), das konkret gelernt wurde bzw. vorliegt. Mithilfe von durch die Semantik spezifizierten Schlussfolgerungsmechanismen kann weiteres *implizites* Wissen abgeleitet werden.

Schauen wir uns die obigen Begriffe am Beispiel der einfachsten Logik, der Aussagenlogik, etwas genauer an. Eine *Proposition* ist eine atomare Aussage, die entweder wahr oder falsch sein kann, wie zum Beispiel $A =$„Der Eiffelturm steht in Paris“ oder $B =$„Kris Kristofferson war ein Musiker“. Eine Menge solcher Propositionen bildet dann die *Signatur* der Sprache und sollte alle atomaren Aussagen enthalten, über die wir in unserer Wissensbasis sprechen möchten. *Logische Formeln* können nun über den Propositionen gebildet werden, um komplexere Sachverhalte und Zusammenhänge zwischen Propositionen darzustellen. Wollen wir beispielsweise den Sachverhalt modellieren, dass wir uns nicht sicher sind, ob Kris Kristofferson ein Musiker (Proposition B) oder ein Schauspieler war (Proposition $C =$„Kris Kristofferson war ein Schauspieler“), aber auf jeden Fall eins von beidem, so repräsentieren wir unser Wissen durch eine *Disjunktion* $\phi_1 = B \vee C$. Hierbei ist ϕ eine zusammengesetzte Formel, die aus den beiden Teilformeln (hier Propositionen) B und C und dem Verknüpfungssymbol $\vee$ besteht. Das Symbol $\vee$ steht hierbei für das umgangssprachliche „Oder“ und die Formel $\phi_1 = B \vee C$ ist genau dann wahr, wenn wenigstens B oder C wahr sind (oder auch beide). Neben der Disjunktion enthält die Sprache der Aussagenlogik noch einige weitere Verknüpfungssymbole. Die *Negation* ($\neg$) einer Formel stellt die Verneinung einer Aussage dar. Beispielsweise steht für $A =$„Der Eiffelturm steht in Paris“ von oben die Formel $\phi_2 = \neg A$ für die Aussage, dass der Eiffelturm *nicht* in Paris steht. Die *Konjunktion* ($\wedge$) verknüpft zwei Formel mit einem logischen „Und“. Für $B =$„Kris Kristofferson war ein Musiker“ und $C =$„Kris Kristofferson war ein Schauspieler“ steht die Formel $\phi_3 = B \wedge C$ für die Aussage, dass Kris Kristofferson sowohl Musiker als auch Schauspieler war (was im übrigen auch der Realität entspricht). Die *Implikation* ($\Rightarrow$) modelliert eine Regel. Mit den Propositionen $D =$„Es regnet“ und $E =$„Das Gras ist nass“ modelliert die Formel $\phi_4 = D \Rightarrow E$ den Sachverhalt, dass, falls es regnet das Gras nass sein wird. Die allgemeine Aussagenlogik erlaubt eine beliebige Verschachtelung der Formeltypen, so dass auch Ausdrücke wie $(A \wedge (B \Rightarrow C)) \vee \neg D$ *wohlgeformt* sind und eine klar definierte Bedeutung haben (mehr dazu unten). Zusammenfassend ist die *Syntax* der Aussagenlogik definiert durch die folgende BNF-Grammatik:

$$\phi ::= a \mid \neg\phi \mid \phi \wedge \phi \mid \phi \vee \phi \mid \phi \Rightarrow \phi$$

wobei a eine Proposition aus einer gegebenen Signatur ist (und Klammerung zur Strukturierung benutzt werden können).

Kommen wir nun zur *Semantik* der Aussagenlogik. Diese beschreibt die *Bedeutung* der Formeln in formaler und maschinenlesbarer Art und Weise und erlaubt die Ableitung eines *Schlussfolgerungsbegriffs*, der es ermöglicht aus einer Menge gegebener Formeln weitere Formeln abzuleiten (siehe unten). Die Semantik der Aussagenlogik definiert man üblicherweise *modelltheoretisch*, d. h., man benutzt Strukturen, die eine vollständige Beschreibung der Welt formalisieren und gibt an, in welche dieser Welten eine gegebene Formel erfüllt ist. Sei Sig eine aussagenlogische Signatur, also die Menge aller Aussagen, über die wir in unserer Logik reden möchten, wie beispielsweise $\mathsf{Sig} = \{A, B, C\}$ von oben (mit $A =$„Der Eiffelturm steht in Paris“, $B =$„Kris Kristofferson war ein Musiker“, $C =$„Kris Kristofferson war ein Schauspieler“). Eine *aussagenlogische Interpretation* (oder *mögliche Welt*) ist eine Funktion $\omega : \mathsf{Sig} \to \{\mathsf{true}, \mathsf{false}\}$, also eine Funktion, die jeder Proposition einen *Wahrheitswert* true (*wahr*) oder false (*falsch*) zuweist, und damit eine mögliche Situation vollständig beschreibt. Die Interpretation ω_1 mit $\omega_1(A) = \mathsf{true}$, $\omega_1(B) = \mathsf{true}$ und $\omega_1(C) = \mathsf{false}$ beschreibt damit die Welt, in der der Eiffelturm in Paris steht ($\omega_1(A) = \mathsf{true}$) und Kris Kristofferson ein Musiker ($\omega_1(B) = \mathsf{true}$), aber kein Schauspieler ($\omega_1(C) = \mathsf{false}$) war. Die Menge aller möglichen Kombinationen von Wahrheitswertzuweisungen spannt damit die Menge aller möglichen Situationen auf.

Wir sagen, dass eine Interpretation ω eine Proposition $a \in \mathsf{Sig}$ *erfüllt*, wenn $\omega(a) = \mathsf{true}$ gilt. Wir schreiben dies auch als $\omega \models a$. Der Erfüllbarkeitsbegriff kann nun auf beliebige Formeln der aussagenlogischen Sprache erweitert werden, indem wir die informellen Bedeutungen von oben nun formal beschreiben. Ist ϕ eine Disjunktion von zwei Teilformeln, also $\phi = \psi_1 \vee \psi_2$ für zwei Formeln ψ_1 und

ψ_2, so erfüllt eine Interpretation ω die Formel ϕ, genau dann wenn ω die Formel ψ_1 erfüllt oder ω die Formel ψ_2 erfüllt (oder beide). Ist ϕ die Negation einer Formel ψ, also $\phi = \neg\psi$, so erfüllt eine Interpretation ω die Formel ϕ, genau dann, wenn ω die Formel ψ nicht erfüllt. Ist ϕ eine Konjunktion von zwei Teilformeln, also $\phi = \psi_1 \wedge \psi_2$ für zwei Formeln ψ_1 und ψ_2, so erfüllt eine Interpretation ω die Formel ϕ, genau dann wenn ω die Formel ψ_1 erfüllt und ω die Formel ψ_2 erfüllt. Ist ϕ eine Implikation $\phi = \psi_1 \Rightarrow \psi_2$, so erfüllt eine Interpretation ω die Formel ϕ, genau dann wenn entweder ω die Formel ψ_1 nicht erfüllt (in diesem Fall ist die Regel nicht anwendbar und damit trivial erfüllt) oder ω erfüllt sowohl ψ_1 also auch ψ_2. Es ist zu beachten, dass trotz ihrer rekursiven Natur diese Spezifikation von *Erfüllbarkeit* wohldefiniert ist, da bei jeder Anwendung einer dieser Fälle, Erfüllbarkeit einer Formel auf Erfüllbarkeit kleinerer Formeln reduziert wird und schließlich die Erfüllbarkeit einer komplexen Formel auf die Wahrheitswerte der Propositionen zurückgeführt wird. Betrachten wir als Beispiel die Formel $\phi = A \wedge (B \vee C)$ und die Interpretation ω_1 mit $\omega_1(A) = \mathsf{true}$, $\omega_1(B) = \mathsf{true}$ und $\omega_1(C) = \mathsf{false}$ von oben (der Eiffelturm steht in Paris, Kris Kristofferson ist ein Musiker, aber kein Schauspieler), dann können wir sehen, dass $\omega_1 \models \phi$ gilt:

1. Es gilt $\omega \models A$, $\omega \models B$ und *nicht* $\omega \models C$.
2. Wegen $\omega \models B$ (und unabhängig von C) gilt $\omega \models B \vee C$.
3. Wegen $\omega \models A$ und $\omega \models B \vee C$ gilt $\omega \models A \wedge (B \vee C)$.

Gilt $\omega \models \phi$ für eine Formel ϕ und eine Interpretation ω, so sagen wir auch, dass ω ein *Modell* von ϕ ist.

Die Semantik spezifiziert die Bedeutung von Formeln in formaler Weise und definiert dadurch, welche Formeln in einer gegebenen Welt wahr sind. In praktischen Szenarien wissen wir aber üblicherweise nicht, in welcher Welt wir uns befinden, wir haben nur *unvollständiges* Wissen über unsere Welt. Beispielsweise könnten wir uns sicher sein, dass der Eiffelturm in Paris steht und das Kris Kristofferson entweder Musiker oder Schauspieler war (oder beides). Dann können wir unser gesamtes Wissen über unsere Welt in der Formel $\phi = A \wedge (B \vee C)$ zusammenfassen. Die Formel ϕ beschreibt dann unsere *Wissensbasis*, also das explizite Wissen, von dessen Wahrheitsgehalt wir überzeugt sind. Die Modelle von ϕ sind die folgenden drei Interpretationen (von denen wir ω_1 oben schon gesehen haben):

$$
\begin{array}{lll}
\omega_1(A) = \mathsf{true} & \omega_1(B) = \mathsf{true} & \omega_1(C) = \mathsf{false} \\
\omega_2(A) = \mathsf{true} & \omega_2(B) = \mathsf{false} & \omega_2(C) = \mathsf{true} \\
\omega_3(A) = \mathsf{true} & \omega_3(B) = \mathsf{true} & \omega_3(C) = \mathsf{true}
\end{array}
$$

Aus der Perspektive von ϕ sind alle drei Interpretationen ω_1, ω_2, und ω_3 *mögliche Welten*. Wenn wir ϕ als wahr annehmen, wissen wir, dass wir in einer dieser drei Welten sein müssen, aber nicht notwendigerweise in welcher.

Wir erweitern unser Beispiel nun noch etwas. Wir wissen beispielsweise, dass berühmte Musiker immer irgendwann auch Schauspieler werden und dass Kris Kristofferson berühmt war. Sei also F eine Proposition mit $F =$„Kris Kristofferson war berühmt“, dann ist unsere neue Wissensbasis ϕ' gegeben durch

$$\phi' = A \wedge (B \vee C) \wedge (B \wedge F \Rightarrow C) \wedge F$$

Die Wissensbasis ϕ' beschreibt immer noch nicht eine eindeutige Welt und in der Praxis sind wir auch selten daran interessiert genau die korrekte Welt zu bestimmen. Wir sind vielmehr daran interessiert, welche *Schlussfolgerungen* wir weiterhin ziehen können, d. h., welche weiteren interessanten Sachverhalte müssen zwangsläufig aus unserer unvollständigen Weltsicht folgen? Wir können beispielsweise sehen, dass wir aus ϕ' schlussfolgern können, dass Kris Kristofferson ein Schauspieler gewesen sein muss:

1. Wegen $(B \vee C)$ war Kris Kristofferson wenigstens ein Musiker oder Schauspieler. War Kris Kristofferson also kein Musiker, so muss er notwendigerweise ein Schauspieler gewesen sein.
2. War Kris Kristofferson aber ein Musiker, so folgt aufgrund der Tatsache, dass er berühmt war (F) und weil berühmte Musiker stets Schauspieler sind $(B \wedge F \Rightarrow C)$, dass er auch Schauspieler gewesen sein muss.

Wir sehen also, dass C eine Schlussfolgerung aus ϕ' ist. Wir schreiben dies üblicherweise auch als $\phi' \models C$. Im Allgemeinen ist der *Schlussfolgerungsbegriff* formal wie folgt definiert: eine Formel ψ *folgt* aus einer Formel ϕ, geschrieben $\phi \models \psi$, wenn jedes Modell von ϕ auch ein Modell von ψ ist. Für unsere Wissensbasis ϕ' sind gerade die folgenden beiden Interpretationen ω'_1 und ω'_2 die (einzigen) Modelle:

$$\begin{array}{llll} \omega'_1(A) = \mathsf{true} & \omega'_1(B) = \mathsf{true} & \omega'_1(C) = \mathsf{true} & \omega'_1(F) = \mathsf{true} \\ \omega'_2(A) = \mathsf{true} & \omega'_2(B) = \mathsf{false} & \omega'_2(C) = \mathsf{true} & \omega'_2(F) = \mathsf{true} \end{array}$$

Es gilt weiterhin offensichtlich $\omega'_1 \models F$ und $\omega'_2 \models F$ und somit $\phi' \models F$.

Ist ϕ unsere Wissensbasis und gilt $\phi \models \psi$ für eine weitere Formel ψ, so nennen wir ψ auch *implizites Wissen* (im Gegensatz zu unserem expliziten Wissen ϕ). Grundsätzlich stellt ϕ also eine Repräsentation unseres Gesamtwissens dar und wir müssen zusätzliche Arbeit leisten, um weiteres (implizites) Wissen daraus abzuleiten. Zu den genauen algorithmischen Herausforderungen für die

Ableitung weiteren Wissens wird in Kapitel 7 genauer eingegangen. Die Art der Schlussfolgerung von ϕ nach ψ nennen wir auch *deduktiven Schluss* oder *Deduktion*. Die Deduktion ist die einzige Form des logisch *korrekten* Schlussfolgerns: gegeben, dass ϕ wahrhaftig gilt, muss bei $\phi \models \psi$ auch ψ notwendigerweise wahrhaftig gelten. Neben der Deduktion gibt es aber auch noch weitere Schlussfolgerungstypen, die zwar nicht notwendigerweise logisch korrekt sind, aber dennoch *plausibel* sein können. Die *Abduktion* ist ein Schlussfolgerungstyp, der gegebenen Beobachtungen auf die Ursache von Beobachtungen schließt. Betrachten wir dazu noch einmal die Propositionen $D =$„Es regnet“ und $E =$„Das Gras ist nass“, sowie die Formel $\phi_4 = D \Rightarrow E$, die die Regel „Wenn es regnet, wird das Gras nass“ modelliert. Stellt ϕ_4 unser gesamtes Wissen dar und wir beobachten, dass das Gras naß ist (also E gilt), so könnten wir daran interessiert sein, den Grund für diese Beobachtung herauszufinden. Ein *abduktiver Schluss* wäre in diesem Fall zu schlussfolgern, dass es geregnet haben muss, da $D \wedge \phi_4 \models E$ gilt. Allgemein formuliert, suchen wir bei der Abduktion bei gegebener Wissensbasis ϕ und Beobachtung ψ eine Formel ϕ', sodass $\phi' \wedge \phi \models \psi$ gilt. Wie man sehen kann, muss dieser Schluss nicht immer korrekt sein. Nehmen wir an, dass es auf dem Gras auch noch einen Rasensprenger gibt (wir haben also eine weitere Proposition $G =$„Der Rasensprenger ist an“) und wir wissen, dass bei eingeschaltetem Rasensprenger das Gras auch nass wird ($G \Rightarrow E$), so ist Regen nicht mehr die einzige mögliche Erklärung, um das nasse Gras zu erklären. Eine weitere plausible Schlussfolgerungsmethode ist die *Induktion*, bei der wir bei Vorhandensein einer Eigenschaften bei (vielen) Einzelfällen auf eine allgemeine Regel schließen. Beobachten wir beispielsweise, dass Amseln fliegen können, dass Raben fliegen können und dass Störche fliegen können, so könnten wir schlussfolgern, dass alle Vögel fliegen können. Auch wenn der induktive Schluss hier plausibel ist (und in vielen anderen Fällen auch angebracht), so sehen wir auch hier, dass er nicht notwendigerweise korrekt sein muss.[1] Schließlich gibt es mit dem Pinguin einen Vogel, der nicht fliegen kann. Wie man besser mit Ausnahmen wie unserem Pinguin umgehen kann wird in Kapitel 8 diskutiert und Kapitel 9 beschäftigt sich weiterhin mit *quantitativen* Schlussfolgerungen, also beispielsweise solche über Wahrscheinlichkeiten. Im nächsten Abschnitt beschäftigen wir uns aber zunächst noch mit weiteren klassischen Methoden und Problemen der *Wissensrepräsentation und -verarbeitung.*

1 Zu beachten ist allerdings, dass die *vollständige Induktion* als mathematisches Werkzeug durchaus eine korrekte Schlussfolgerung darstellt.

3 Methoden und Algorithmen

Wissen sollte so repräsentiert werden, dass es effektiv genutzt und verändert werden kann. Um die richtigen Wissensverarbeitungsmechanismen bereitstellen zu können, ist die *Natur* oder der *Typ* des Wissens von Relevanz, d. h., besteht das Wissen aus einfachen aussagenlogischen Fakten, aus relationalen Beziehungen, gibt es quantitative Gewichtungen des Wissens, oder behandelt das Wissen verschiedene Zeitpunkte? Generische Systeme für die Behandlung von Wissen, wie z. B. klassische Datenbanksysteme, erlauben üblicherweise nur die Modellierung mit sehr klassischen Semantiken (wie die der Aussagenlogik). Die Semantik bestimmt jedoch auch welche spezifischen Methoden der Wissensverarbeitung eingesetzt werden können.

Im einfachsten Fall besteht ein *Wissensrepräsentationssystem* aus einer Wissensbasis K und einer Möglichkeit, Anfragen ($\models$) an das Wissen zu stellen:

$$\langle K, \models \rangle$$

Hierbei kann K beispielsweise eine Menge aussagenlogischer Formeln und $\models$ die aussagenlogische Schlussfolgerungsrelation aus dem vorherigen Abschnitt darstellen. Dies stellt allerdings nur eine sehr einfache Realisierung eines Wissensrepräsentationssystems dar. Im Folgenden werden wir mit der Beschreibungslogik $\mathcal{ALC}$ und mit Modallogiken zwei weitere, anspruchsvollere, Möglichkeiten der Wissensrepräsentation kennenlernen. Danach werden wir auf die Veränderung von Wissensrepräsentation anhand von neuen Informationen eingehen.

3.1 Beschreibungslogik

Als Beispiel für einen Wissensrepräsentationsformalismus werden wir die Beschreibungslogik $\mathcal{ALC}$ (*Attributive Concept Description Language with Complements*; Deutsch: „Attributive Begriffsbeschreibungssprache mit Komplementen“) [25] betrachten. Beschreibungslogiken konzentrieren sich auf die Beschreibung von Beziehungen zwischen *Begriffen* und *Individuen*. Die Wissensbasen von Beschreibungslogiken sind geteilt in eine Komponente, die Wissen über Individuen enthält (die *ABox*), und eine Komponente (die *TBox*), die Wissen über Beziehungen zwischen Konzepten enthält (terminologisches Wissen).

3.1.1 Konzepte und deren Interpretation

Wir beginnen damit Konzeptbeschreibungen zu definieren und wie wir diese interpretieren. Ein Konzept C wird als die Menge aller Individuen verstanden, die zu dem Konzept gehören. Beispielsweise beschreibt das Konzept Koala die Menge aller konkreten Lebewesen, die Koalas sind.

Grundbausteine der Sprache sind *Konzeptsymbole*, oder auch *Konzeptnamen* genannt, die für atomare Konzepte stehen und *Rollensymbole*, die für (atomare) Relationen zwischen Individuen stehen. Die Vereinbarung ist dabei, dass Konzept- und Rollensymbole stets mit einem Großbuchstaben beginnen. Es folgt die rekursiv definierte formale Sprache für eine komplexe Konzeptbeschreibung C, wobei A für ein beliebiges atomares Konzeptsymbol steht und R für ein beliebiges atomares Rollensymbol:

$$
\begin{array}{lcll}
C & \longrightarrow & \top & \text{(universelles Konzept)} \\
 & \mid & \bot & \text{(leeres Konzept)} \\
 & \mid & A & \text{(atomares Konzept)} \\
 & \mid & C \sqcap C & \text{(Konzeptkonjunktion)} \\
 & \mid & C \sqcup C & \text{(Konzeptdisjunktion)} \\
 & \mid & \neg C & \text{(Konzeptnegation)} \\
 & \mid & \forall R.C & \text{(universelle Restriktion)} \\
 & \mid & \exists R.C & \text{(existenzielle Restriktion)}
\end{array}
$$

Wie bereits eingangs erwähnt werden wir Konzepte als Mengen von Individuen verstehen. Es ist daher sinnvoll, Operationen zu definieren, die zu Operationen und Objekten aus der Mengentheorie korrespondieren. Aus dieser Perspektive ist es natürlich, das *universelle Konzept* (das alle Individuen umfasst) und das *leere Konzept* (das kein Individuum enthält) zu betrachten. Ebenso sind die *Vereinigung von Konzepten* $C \sqcup D$ und der *Schnitt von Konzepten* $C \sqcap D$ natürliche Operationen. *Konzeptnegation* $\neg C$ steht für das Konzept, das alle Individuen hält, die nicht im Konzept C sind. Die *universelle Restriktion* $\forall R.C$ beschreibt ein Konzept, das (via R) nur in Rollenbeziehungen zu Individuen aus C steht. Beispielsweise steht $\forall$Isst.Pflanze für das Konzept, das alle Individuen enthält, die nur Pflanzen essen. Die *existenzielle Restriktion* $\exists R.C$ beschreibt ein Konzept, das (via R) in Beziehungen zu mindestens einem Individuum aus C steht. Beispielsweise steht $\exists$Isst.Pflanze für das Konzept, dass alle Individuen enthält, die mindestens eine Pflanze essen (aber eventuell auch noch etwas anderes essen).

Man beachte, dass wir oben formal von *Konzeptsymbolen* und *Rollensymbolen* und nicht direkt von *Konzepten* und *Rollen* reden. In der Beschreibungslogik unterscheiden wir zwischen den Symbolen und der Bedeutung, die wir diesen

$$
\begin{aligned}
\Delta^{\mathcal{I}_{\mathsf{Tiere}}} &= \{\text{klaus}, \text{bernd}, \text{leo}, \text{eukalyptus}, \text{wurzel}, \text{fleisch}\} \\
\mathsf{klaus}^{\mathcal{I}_{\mathsf{Tiere}}} &= \text{klaus} & \mathsf{eukalyptus}^{\mathcal{I}_{\mathsf{Tiere}}} &= \text{eukalyptus} \\
\mathsf{bernd}^{\mathcal{I}_{\mathsf{Tiere}}} &= \text{bernd} & \mathsf{wurzel}^{\mathcal{I}_{\mathsf{Tiere}}} &= \text{wurzel} \\
\mathsf{leo}^{\mathcal{I}_{\mathsf{Tiere}}} &= \text{leo} & \mathsf{fleisch}^{\mathcal{I}_{\mathsf{Tiere}}} &= \text{fleisch} \\
\mathsf{Koala}^{\mathcal{I}_{\mathsf{Tiere}}} &= \{\text{klaus}\} & \mathsf{Bär}^{\mathcal{I}_{\mathsf{Tiere}}} &= \{\text{bernd}\} \\
\mathsf{Löwe}^{\mathcal{I}_{\mathsf{Tiere}}} &= \{\text{leo}\} & \mathsf{Pflanze}^{\mathcal{I}_{\mathsf{Tiere}}} &= \{\text{eukalyptus},\ \text{wurzel}\} \\
\mathsf{Isst}^{\mathcal{I}_{\mathsf{Tiere}}} &= \{(\text{klaus}, \text{eukalyptus}),\ (\text{bernd}, \text{wurzel}),\ (\text{bernd}, \text{fleisch}),\ (\text{leo}, \text{fleisch})\}
\end{aligned}
$$

Abbildung 1: Beispielinterpretation $\mathcal{I}_{\mathsf{Tiere}}$.

zuweisen. Neben Konzeptsymbolen und Rollensymbolen werden wir später auch über konkrete Individuennamen reden. Die Vereinbarung ist dabei, dass Individuennamen stets mit Kleinbuchstaben beginnen.

Wir definieren nun die formale Semantik von $\mathcal{ALC}$. Eine (beschreibungslogische) Interpretation ist ein Paar

$$\mathcal{I} = \langle \Delta^{\mathcal{I}}, \cdot^{\mathcal{I}} \rangle \ .$$

Die erste Komponente $\Delta^{\mathcal{I}}$ ist die *Domäne* von $\mathcal{I}$ und enthält die konkreten *Individuen*. Konzeptsymbole, Rollensymbole und Individuennamen werden durch $\cdot^{\mathcal{I}}$ über $\Delta^{\mathcal{I}}$ interpretiert. Das heißt, dass jedem Konzeptsymbol C eine Teilmenge $C^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}}$ zugewiesen wird, jedem Rollensymbol R eine Teilmenge $R^{\mathcal{I}} \subseteq \Delta^{\mathcal{I}} \times \Delta^{\mathcal{I}}$ zugewiesen wird, und jedem Individuennamen a wird ein Individuum $a^{\mathcal{I}}$ aus $\Delta^{\mathcal{I}}$ zugewiesen. Wir verlangen zusätzlich, dass jeder Individuennamen einzigartig interpretiert wird. Das heißt, für zwei unterschiedliche Individuennamen a und b gilt $a^{\mathcal{I}} \neq b^{\mathcal{I}}$. Dies wird auch als die *Unique Name Assumption* bezeichnet.

Eine beispielhafte Interpretation $\mathcal{I}_{\mathsf{Tiere}}$ ist in Abbildung 1 angeben. Diese Interpretation modelliert einen kleinen Zoo mit Tieren und deren Essgewohnheiten. Hier gibt es zwei Arten von Individuen: bestimmte Tiere („klaus", „bernd", „leo"), und Gegenstände, die diese eventuell essen („eukalyptus", „wurzel", „fleisch"). Die atomaren Konzepte in $\mathcal{I}_{\mathsf{Tiere}}$ sind Koala, Löwe, Bär, Pflanze. Die einzige Rolle ist Isst. Man beachte, dass es sich beispielsweise bei „klaus" um das Individuum in der Domäne $\Delta^{\mathcal{I}_{\mathsf{Tiere}}}$ handelt, während klaus ein Individuennamen ist, der als „klaus" interpretiert wird.

In der formalen Semantik werden komplexe Konzeptbeschreibungen unter einer Interpretation selber zu einem Konzept ausgewertet. Gegeben eine Interpretation $\mathcal{I} = \langle \Delta^{\mathcal{I}}, \cdot^{\mathcal{I}} \rangle$, ist die Semantik wie folgt definiert:

$$\begin{aligned}
\top^{\mathcal{I}} &= \Delta^{\mathcal{I}} \\
\bot^{\mathcal{I}} &= \emptyset \\
(C \sqcap D)^{\mathcal{I}} &= C^{\mathcal{I}} \cap D^{\mathcal{I}} \\
(C \sqcup D)^{\mathcal{I}} &= C^{\mathcal{I}} \cup D^{\mathcal{I}} \\
(\neg C)^{\mathcal{I}} &= \Delta^{\mathcal{I}} \setminus C^{\mathcal{I}} \\
(\forall R.C)^{\mathcal{I}} &= \{a \in \Delta^{\mathcal{I}} \mid \text{für alle } b \text{ mit } (a,b) \in R^{\mathcal{I}} \text{ gilt } b \in C^{\mathcal{I}}\} \\
(\exists R.C)^{\mathcal{I}} &= \{a \in \Delta^{\mathcal{I}} \mid \text{es gibt ein } b \text{ mit } (a,b) \in R^{\mathcal{I}} \text{ und } b \in C^{\mathcal{I}}\}
\end{aligned}$$

Wir sehen, dass die Semantik sich sehr verträglich mit der Interpretation als Mengen verhält. Das Konzeptsymbol $\top$ wird als die Menge aller Individuen interpretiert, und $\bot$ wird stets als das leere Konzept. Eine Konzeptkonjunktion $C \sqcap D$ wird als die Menge aller Individuen verstanden, die in beiden Konzepten C und D vorkommen, und, analog dazu, wird eine Konzeptdisjunktion $C \sqcup D$ als die Vereinigung aller Individuen in C und D interpretiert. $\neg C$ wird als die Menge aller Individuen interpretiert, die nicht in C vorkommen[2] (bezogen auf die Domäne $\Delta^{\mathcal{I}}$). Universelle Restriktionen $\forall R.C$ werden als die Menge alle Individuen interpretiert, die über R nur mit Individuen aus C in Beziehung stehen. Existenzielle Restriktionen $\exists R.C$ werden als die Menge alle Individuen interpretiert, die über R mit mindestens einem Individuum in C in Beziehung stehen.

Als konkretes Beispiel betrachten wir die Interpretation $\mathcal{I}_{\mathsf{Tiere}}$ aus Abbildung 1. Hier gilt $\mathsf{Pflanze}^{\mathcal{I}_{\mathsf{Tiere}}} = \{\text{eukalyptus}, \text{wurzel}\}$ und die Rolle Isst wird wie folgt interpretiert:

$$\mathsf{Isst}^{\mathcal{I}_{\mathsf{Tiere}}} = \{(\text{klaus}, \text{eukalyptus}),\ (\text{bernd}, \text{wurzel}),\ (\text{bernd}, \text{fleisch}),\ (\text{leo}, \text{fleisch})\}$$

Unter der oben definierten Semantik gilt $(\forall \mathsf{Isst}.\mathsf{Pflanze})^{\mathcal{I}_{\mathsf{Tiere}}} = \{\text{klaus}\}$. Dies ist der Fall, da für alle Paare $(a,b) \in \mathsf{Isst}^{\mathcal{I}_{\mathsf{Tiere}}}$ mit $a = \text{klaus}$ gilt, dass b ein Element in $\mathsf{Pflanze}^{\mathcal{I}_{\mathsf{Tiere}}}$ ist. Hingegen gilt $\text{bernd} \notin (\forall \mathsf{Isst}.\mathsf{Pflanze})^{\mathcal{I}_{\mathsf{Tiere}}}$, da das Individuum bernd über Isst mit einem Element in Beziehung steht, das kein Element in $\mathsf{Pflanze}$ ist. Ein Beispiel dafür ist das Paar $(\text{bernd}, \text{fleisch})$ aus $\mathsf{Isst}^{\mathcal{I}_{\mathsf{Tiere}}}$, wobei fleisch kein Element in $\mathsf{Pflanze}^{\mathcal{I}_{\mathsf{Tiere}}}$ ist. Des Weiteren gilt $(\exists \mathsf{Isst}.\mathsf{Pflanze})^{\mathcal{I}_{\mathsf{Tiere}}} = \{\text{klaus}, \text{bernd}\}$. Insbesondere gilt $\text{bernd} \in (\exists \mathsf{Isst}.\mathsf{Pflanze})^{\mathcal{I}_{\mathsf{Tiere}}}$, da es mit $(\text{bernd}, \text{wurzel})$ ein Paar in $\mathsf{Isst}^{\mathcal{I}_{\mathsf{Tiere}}}$ gibt, das bernd in Beziehung mit einem Element aus $\mathsf{Pflanze}^{\mathcal{I}_{\mathsf{Tiere}}}$ setzt.

2 Die Negation verhält sich also wie ein Komplement auf der semantischen Ebene. Das $\mathcal{C}$ in $\mathcal{ALC}$ steht auch für Komplement (Englisch: „*complement*“).

$$\begin{array}{ll} \mathsf{Koala}(\mathsf{klaus}) & \mathsf{Isst}(\mathsf{klaus}, \mathsf{eukalyptus}) \\ \mathsf{Bär}(\mathsf{bernd}) & \mathsf{Isst}(\mathsf{bernd}, \mathsf{wurzeln}) \\ \mathsf{Löwe}(\mathsf{leo}) & \mathsf{Isst}(\mathsf{bernd}, \mathsf{fleisch}) \\ \mathsf{Pflanze}(\mathsf{eukalyptus}) & \mathsf{Isst}(\mathsf{leo}, \mathsf{fleisch}) \\ \mathsf{Pflanze}(\mathsf{wurzeln}) & \end{array}$$

Abbildung 2: Beispiel ABox $\mathcal{A}_{\mathsf{Tiere}}$.

3.1.2 Wissen über Individuen

Die Beschreibungslogik $\mathcal{ALC}$ erlaubt zwei Arten von expliziten Aussagen über Individuen. Die erste sind Aussagen über die Zugehörigkeit zu einem Konzept. Beispielsweise repräsentieren wir die Aussage, dass das Individuum „Klaus“ ein Koala ist, mit $\mathsf{Koala}(\mathsf{klaus})$. Die zweite Art von Aussagen sind Rollenbeziehungen zwischen zwei Individuen. Beispielsweise könnten wir die Aussage, dass „Klaus“ Eukalyptus isst, formal durch $\mathsf{Isst}(\mathsf{klaus}, \mathsf{eukalyptus})$ repräsentieren.

Die allgemeine Syntax für eine Individuenaussage A ist im Folgenden angegeben, wobei C eine komplexe Konzeptbeschreibung ist, R ein Rollensymbol und a, b für Individuennamen stehen:

$$\begin{array}{cccl} A & \longrightarrow & C(a) & \text{(Konzeptzuweisung)} \\ & \mid & R(a,b) & \text{(Rollenbeschreibung)} \end{array}$$

Mengen von Konzeptzuweisungen und Rollenbeschreibungen werden als *ABox* bezeichnet. Der Name *ABox* leitet sich vom englischen Wort *assertion* (Deutsch: *Behauptung* oder auch *Zusicherung*) ab. Eine beispielhafte ABox $\mathcal{A}_{\mathsf{Tiere}}$ für unser Beispiel mit einem Koala „Klaus“, einigen anderen Tieren und deren Ernährungsbedürfnissen ist in Abbildung 2 dargestellt.

Wir sagen, dass eine Interpretation $\mathcal{I} = \langle \Delta^{\mathcal{I}}, \cdot^{\mathcal{I}} \rangle$ eine Konzeptzuweisung $A(a)$ *erfüllt*, geschrieben $\mathcal{I} \models A(a)$, genau dann wenn $a^{\mathcal{I}} \in A^{\mathcal{I}}$ gilt. Eine Interpretation $\mathcal{I} = \langle \Delta^{\mathcal{I}}, \cdot^{\mathcal{I}} \rangle$ *erfüllt* eine Rollenzuweisung $R(a, b)$, geschrieben $\mathcal{I} \models R(a, b)$, genau dann wenn $(a^{\mathcal{I}}, b^{\mathcal{I}}) \in R^{\mathcal{I}}$ gilt. Die Interpretation $\mathcal{I}$ erfüllt die ABox $\mathcal{A}$, geschrieben $\mathcal{I} \models \mathcal{A}$, genau dann wenn $\mathcal{I} \models F$ für alle $F \in \mathcal{A}$ gilt.

Für unser Beispiel können wir zeigen, dass die Interpretation $\mathcal{I}_{\mathsf{Tiere}}$ die ABox $\mathcal{A}_{\mathsf{Tiere}}$ erfüllt. Schnell lässt sich sehen, dass alle Konzeptzuweisungen in $\mathcal{A}_{\mathsf{Tiere}}$ erfüllt sind von $\mathcal{I}_{\mathsf{Tiere}}$. So gilt beispielsweise, dass die Konzeptzuweisungen $\mathsf{Koala}(\mathsf{klaus})$ erfüllt ist, da $\mathsf{klaus}^{\mathcal{I}_{\mathsf{Tiere}}} \in \mathsf{Koala}^{\mathcal{I}_{\mathsf{Tiere}}}$ gilt. Die beiden Rollenbeschreibungen $\mathsf{Isst}(\mathsf{klaus}, \mathsf{eukalyptus})$ und $\mathsf{Isst}(\mathsf{leo}, \mathsf{fleisch})$ sind auch erfüllt. Erstere ist erfüllt, da $(\text{klaus}, \text{eukalyptus}) \in \mathsf{Isst}^{\mathcal{I}_{\mathsf{Tiere}}}$ gilt, und die zweite ist erfüllt, da $(\text{leo}, \text{fleisch}) \in \mathsf{Isst}^{\mathcal{I}_{\mathsf{Tiere}}}$ gilt.

3.1.3 Terminologisches Wissen

Die ABox bietet die Möglichkeit Faktenwissen über konkrete einzelne Individuen und deren direkte Beziehung explizit auszudrücken. Neben solchem faktischen Wissen gibt es aber auch generisches Wissen. In Beschreibungslogiken lässt sich insbesondere der Zusammenhang zwischen Konzepten formalisieren. Dieses generische Wissen wird in der TBox hinterlegt.

Ausdrücke, die Konzepte in Beziehung zueinander setzen werden als *terminologische Axiome* bezeichnet. Die terminologischen Axiome, die wir hier betrachten, sind von der Form

$$C \sqsubseteq D \quad \text{oder} \quad C \equiv D$$

wobei C, D komplexe Konzepte sind. Axiome der Form $C \sqsubseteq D$ werden *Inklusionen* genannt und bedeuten informell, dass das Konzept C ein Teil des Konzepts D ist. Axiome der Form $C \equiv D$ werden als *Äquivalenzen* bezeichnet, und drücken aus, dass C und D dieselben Individuen beschreiben. Eine Menge von terminologischen Axiomen $\mathcal{T}$ wird als *TBox* bezeichnet.

Die bereits eingeführten Teile der Semantik von $\mathcal{ALC}$ erweitert sich zu TBoxen auf natürliche Art und Weise. Für eine Interpretation $\mathcal{I} = \langle \Delta^{\mathcal{I}}, \cdot^{\mathcal{I}} \rangle$ und komplexe Konzeptbeschreibungen definieren wir:

$$\begin{aligned} \mathcal{I} \models C \sqsubseteq D &\quad \text{genau dann wenn} \quad C^{\mathcal{I}} \subseteq D^{\mathcal{I}} \\ \mathcal{I} \models C \equiv D &\quad \text{genau dann wenn} \quad C^{\mathcal{I}} = D^{\mathcal{I}} \end{aligned}$$

Wir sagen $\mathcal{I}$ erfüllt eine TBox $\mathcal{T}$, geschrieben $\mathcal{I} \models \mathcal{T}$ genau dann wenn $\mathcal{I}$ jedes terminologisches Axiom in $\mathcal{T}$ erfüllt.

Insbesondere Äquivalenzen können wir nutzen, um Konzepte mithilfe anderer Konzepte zu definieren. Wir können beispielsweise die Konzepte der Pflanzenfresser, Allesfresser und Fleischfresser in einer TBox direkt definieren:

$$\begin{aligned} \mathsf{Herbivoren} &\equiv \forall \mathsf{Isst}.\mathsf{Pflanze} \\ \mathsf{Omnivoren} &\equiv \exists \mathsf{Isst}.\mathsf{Pflanze} \sqcap \exists \mathsf{Isst}.\neg \mathsf{Pflanze} \\ \mathsf{Canivoren} &\equiv \forall \mathsf{Isst}.\neg \mathsf{Pflanze} \end{aligned}$$

Wir erweitern das Beispiel noch durch die Inklusionen

$$\mathsf{Koala} \sqsubseteq \mathsf{Herbivoren} \qquad \mathsf{Bär} \sqsubseteq \mathsf{Omnivoren} \qquad \mathsf{Löwe} \sqsubseteq \mathsf{Canivoren}$$

die beschreiben, dass Koalas Pflanzenfresser, Bären Allesfresser und Löwen Fleischfresser sind. Sei $\mathcal{T}_{\mathsf{Tiere}}$ die TBox, die die obigen Äquivalenzen und Inklusionen enthält. Die Interpretation $\mathcal{I}_{\mathsf{Tiere}}$ ist in der Tat ein Modell von $\mathcal{T}_{\mathsf{Tiere}}$. So gilt unter der Interpretation $\mathcal{I}_{\mathsf{Tiere}}$ beispielsweise, dass klaus $\in \mathsf{Herbivoren}^{\mathcal{I}_{\mathsf{Tiere}}}$ gilt. Da eben auch $\mathsf{Koala}^{\mathcal{I}_{\mathsf{Tiere}}} = \{\text{klaus}\}$ gilt, ist auch die Inklusion $\mathsf{Koala} \sqsubseteq \mathsf{Herbivoren}$ erfüllt unter $\mathcal{I}_{\mathsf{Tiere}}$.

3.1.4 Wissensbasen, Anfragen und Komplexität

In den letzten beiden Abschnitten haben wir ABoxen und TBoxen, sowie die Semantik dieser Komponenten definiert. Eine $\mathcal{ALC}$ Wissensbasis ist nun ein Paar $K = \langle \mathcal{A}, \mathcal{T} \rangle$, wobei $\mathcal{A}$ eine ABox ist und $\mathcal{T}$ eine TBox. Die Modelle von K sind alle Interpretationen $\mathcal{I} = \langle \Delta^{\mathcal{I}}, \cdot^{\mathcal{I}} \rangle$, die sowohl $\mathcal{A}$ erfüllen als auch $\mathcal{T}$ erfüllen. Wenn $\mathcal{I}$ ein Modell von K ist, dann schreiben wir auch $\mathcal{I} \models K$. Beispielsweise gilt $\mathcal{I}_{\mathsf{Tiere}} \models \langle \mathcal{A}_{\mathsf{Tiere}}, \mathcal{T}_{\mathsf{Tiere}} \rangle$, denn wie wir bereits zuvor gesehen haben, gilt $\mathcal{I}_{\mathsf{Tiere}} \models \mathcal{A}_{\mathsf{Tiere}}$ und $\mathcal{I}_{\mathsf{Tiere}} \models \mathcal{T}_{\mathsf{Tiere}}$.

Für ein Konzeptbeschreibung C sagen wir, dass C erfüllbar in K ist, falls es ein Modell $\mathcal{I}$ von K gibt, indem $C^{\mathcal{I}}$ nichtleer ist Wir sagen, zwei Konzeptbeschreibungen C, D sind disjunkt bezüglich K, falls $C^{\mathcal{I}} \cap D^{\mathcal{I}} = \emptyset$ für jedes Modell $\mathcal{I}$ von K gilt. Formal können wir in Beschreibungslogiken die folgenden Entscheidungsprobleme betrachten:

Subsumption
Eingabe: Eine $\mathcal{ALC}$ Wissensbasis $K = \langle \mathcal{A}, \mathcal{T} \rangle$ und zwei Konzeptsymbole C, D.
Frage: Gilt $K \models C \sqsubseteq D$?

Konzept-Erfüllbarkeit
Eingabe: Eine $\mathcal{ALC}$ Wissensbasis $K = \langle \mathcal{A}, \mathcal{T} \rangle$ und ein Konzeptsymbol C.
Frage: Ist C erfüllbar in K?

Äquivalenz
Eingabe: Eine $\mathcal{ALC}$ Wissensbasis $K = \langle \mathcal{A}, \mathcal{T} \rangle$ und zwei Konzeptsymbole C, D.
Frage: Gilt $K \models C \equiv D$?

Disjunktheit
Eingabe: Eine $\mathcal{ALC}$ Wissensbasis $K = \langle \mathcal{A}, \mathcal{T} \rangle$ und zwei Konzeptsymbole C, D.
Frage: Sind C, D disjunkt in K?

Aus komplexitätstheoretischer Perspektive lässt sich zeigen, dass sich alle obigen Entscheidungsprobleme auf das Subsumption-Problem zurückführen lassen, denn es gelten folgenden Beziehungen:

$K \models C \equiv D$	genau dann, wenn	$K \models C \sqsubseteq D$ und $K \models D \sqsubseteq C$
C, D sind disjunkt in K	genau dann, wenn	$K \models C \sqcap D \sqsubseteq \bot$
C ist unerfüllbar in K	genau dann, wenn	$K \models C \sqsubseteq \bot$

Im Allgemeinen ist das SUBSUMPTION-Problem für $\mathcal{ALC}$ PSPACE-vollständig [25]. Das heißt unter anderem, es gibt einen Algorithmus, der das SUBSUMPTION-Problem löst und dabei polynomiell viel Speicherplatz benötigt (bezogen auf K, C und D).

Die Beschreibungslogik $\mathcal{ALC}$ ist aus logischer Sicht eine prototypische Beschreibungslogik, da sie die üblichen Verknüpfungen Konjunktion, Disjunktion und Negation hat, die wir bereits aus der Aussagenlogik kennen. Darüber hinaus sind viele Erweiterungen und Einschränkungen von $\mathcal{ALC}$ bekannt. So kann man $\mathcal{ALC}$ um Konstrukte erweitern, die es erlauben zu zählen, wie viele Individuen zu einem Konzept gehören, beziehungsweise die Anzahl der Elemente in Konzepten zu vergleichen. Auch kennt man für diese Erweiterungen die Komplexität der obigen Entscheidungsprobleme sehr genau. Eine gute Quelle für weitere Informationen ist das *Description Logic Handbook* [3]. In der Praxis spielen die verschiedenen Beschreibungslogiken eine große Rolle im Bereich der Ontologien. Namentlich sei hier die standardisierte *Web Ontology Language*[3] (OWL) zu nennen, die auf Beschreibungslogiken basiert und im Ontologie-Editor *Protégé*[4] verwendet wird.

3.2 Wissenslogiken

In diesem Kapitel thematisieren wir Wissenslogiken, die es erlauben, Aussagen darüber zu treffen, welches Wissen ein Agent[5] besitzt. Dies ist hilfreich um zu verstehen, wie man Metainformationen über Wissen repräsentieren kann und wie sich diese verarbeiten lassen.

Ein Katastrophenschutzsystem kann beispielsweise mithilfe solch einer Logik repräsentieren, dass ein Feuerwehrmann nur bestimmte Informationen besitzt.

3 https://www.w3.org/TR/owl-guide/

4 https://protege.stanford.edu/

5 Der Begriff des Agenten wird üblicherweise sehr breit verstanden. Man meint damit ein selbständig agierendes System, wie beispielsweise einen Menschen, Tier, Roboter oder ein Computerprogramm. Das Kapitel 20 über Multiagentensysteme geht genauer auf den Begriff des Agenten ein.

Weiß das System, das der Feuerwehrmann über einen Sammelpunkt informiert ist, so kann das System schlussfolgern, dass der Feuerwehrmann einen Verletzten zu diesem Sammelpunkt bringen wird. Basierend auf dieser Schlussfolgerung könnten das Katastrophenschutzsystem präventiv einen Krankenwagen zu dem Sammelpunkt senden. In dem Fall, in dem das System weiß, dass der Feuerwehrmann nicht über den Sammelpunkt informiert ist, kann das System schlussfolgern, dass ein Feuerwehrmann einen Verletzten zu einer anderen Stelle bringen wird. Basierend auf dieser Schlussfolgerung kann das System anderes reagieren, und einen Krankenwagen zu der anderen Stelle senden. Die Logik erlaubt es also auszudrücken und zu verstehen, welche Schlussfolgerungen der Feuerwehrmann auf Basis seiner (subjektiven) Informationen ziehen kann. Basierend auf dieser Information kann dann entsprechend reagiert werden.

Konkret betrachten wir hier Modallogiken. In diesen Logiken lassen sich Aussagen der Art „der Agent ist überzeugt, dass φ gilt“ und der Art „der Agent hält es für möglich, dass φ gilt“ explizit ausdrücken. Im folgenden Abschnitt definieren wir zunächst die Syntax einer Modallogik und darauf folgend die Standardsemantik.

3.2.1 Syntax der Modallogik

Wir definieren eine formale Sprache, die es erlaubt Aussagen über Wissen zu formulieren. Die Syntax dieser Sprache ist eine Erweiterung der Aussagenlogik von oben:

$$\phi ::= a \mid \neg\phi \mid \phi \wedge \phi \mid \phi \vee \phi \mid \phi \Rightarrow \phi \mid \Box\phi$$

wobei $a \in \mathsf{Sig}$ wie zuvor ein Element der Signatur ist. Zusätzlich definieren wir $\Diamond\varphi$ als Abkürzung für $\neg\Box\neg\varphi$. Man beachte, dass wir auch Verschachtelungen der Art $\Box\Diamond\varphi$ erlauben. Die Operatoren $\Box$ und $\Diamond$ werden auch als Modaloperatoren bezeichnet (abgeleitet von dem sprachwissenschaftlichen Begriff der Modalität). Logiken die Modaloperatoren haben, werden auch als Modallogik bezeichnet.

Die neue Formel $\Box\varphi$ hat die intendierte Bedeutung von „der Agent ist überzeugt, dass φ gilt“. Wir werden weiter unten sehen, dass zu der Regel $\Diamond\varphi$ die Bedeutung „der Agent hält es für möglich, dass φ gilt“ sehr gut passt. Die Formel $\varphi \Rightarrow \Box\varphi$ drückt weiterhin beispielsweise aus, dass der Agent von φ überzeugt ist, wenn φ gilt. Die Formel $\Box\Diamond\varphi$ lässt als „der Agent ist überzeugt, das er φ für möglich hält“ interpretieren.

3.2.2 Semantik der Modallogik

Wir erinnern uns, dass eine aussagenlogische Interpretation $\omega : \mathsf{Sig} \to \{\mathsf{true}, \mathsf{false}\}$ jedem Signaturelement in Sig einen Wahrheitswert zuweist. Eine aussagenlogische Interpretation ist also eine vollständige Beschreibung wie die Welt (bezogen auf die Elemente in Sig) aussehen könnten. Daher werden aussagenlogische Interpretation im Kontext der Aussagenlogik auch als mögliche Welten bezeichnet.

Unsere modallogische Sprache ist reichhaltiger als die Sprache der aussagenlogischen Formeln, da wir mit Formeln wie $\Box\varphi$ in der Lage sind auszudrücken, dass der Agent von φ überzeugt ist. Die Information, woran ein Agent glaubt, ist aber nicht durch aussagenlogische Interpretationen bestimmbar. Daher nutzt man eine Verallgemeinerung der Menge der aussagenlogischen Interpretationen, die mehr Struktur enthält als im Falle der Aussagenlogik. Eine Kripke-Struktur, benannt nach Saul Kripke [18], ist ein Tripel $K = \langle W, R, V\rangle$, wobei

- W eine Menge ist, deren Elemente als (mögliche) Welten bezeichnet werden,
- $R \subseteq W \times W$ eine Relation zwischen W ist, die als *Erreichbarkeitsrelation* bezeichnet wird, und
- $V : W \to \mathcal{P}(\mathsf{Sig})$ ist eine Funktion, die jeder Welt eine Menge von atomaren Aussagen zuweist.

Die Elemente in W haben, zusammen mit V, die Funktion, Aussagen in Sig Wahrheitswerte zuzuweisen. Der formale Unterschied zu aussagenlogischen Interpretationen ist, dass die Elemente in W keine Funktionen sind. Stattdessen beschreibt die Menge $V(w)$ für jede Welt $w \in W$, welche Aussagen in w als wahr interpretiert werden. Alle Aussagen aus Sig, die nicht in $V(w)$ sind, gelten implizit als falsch. Aus R erhalten wir zu jeder Welt $w \in W$ die Menge der von w aus *erreichbaren* Welten. Wir definieren $R(w) := \{w' \in W \mid (w, w') \in R\}$ als die Menge derjenigen Welten w' für die $(w, w') \in R$ gilt. Informell beschreibt $R(w)$ diejenigen Welten, die der Agent für möglich hält, wenn w die „richtige“ Welt ist.

Als Beispiel sei $\mathsf{Sig} = \{P, Q\}$ und die Kripke-Struktur $K_{\mathrm{pq}} = \langle W_{\mathrm{pq}}, R_{\mathrm{pq}}, V_{\mathrm{pq}}\rangle$ mit $W_{\mathrm{pq}} = \{w_1, w_2, w_3, w_4\}$ und

$$\begin{aligned} R_{\mathrm{pq}} = \{\ & (w_1, w_1), (w_1, w_2), (w_3, w_3), & V_{\mathrm{pq}}(w_1) &= \{P, Q\} & V_{\mathrm{pq}}(w_3) &= \{Q\} \\ & (w_2, w_2), (w_2, w_3), (w_2, w_4)\ \} & V_{\mathrm{pq}}(w_2) &= \{P\} & V_{\mathrm{pq}}(w_4) &= \emptyset \end{aligned}$$

gegeben. Betrachten wir beispielsweise die Welt w_3 aus K_{pq}. Für diese Welt gilt $V_{\mathrm{pq}}(w_3) = \{Q\}$. Dies bedeutet, dass Q den Wahrheitswert wahr in w_3 erhält und P den Wahrheitswert falsch in w_3. Wenn wir die Erreichbarkeitsrelation R_{pq}

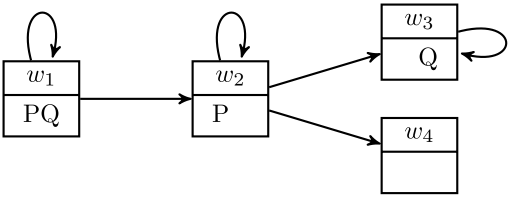

Abbildung 3: Graph-Darstellung der Kripke-Struktur K_{pq}.

betrachten, erhalten wir:

$$R_{\text{pq}}(w_1) = \{w_1, w_2\} \qquad R_{\text{pq}}(w_2) = \{w_2, w_3, w_4\}$$
$$R_{\text{pq}}(w_3) = \{w_3\} \qquad R_{\text{pq}}(w_4) = \emptyset\}$$

Abbildung 3 stellt K_{pq} graphisch dar.

Die Bedeutung der Erreichbarkeitsrelation wird insbesondere klar, wenn wir die Semantik definieren. Wenn $K = \langle W, R, V \rangle$ eine Kripke-Struktur ist, $w \in W$ eine Welt aus W ist, und φ eine modallogische Formel ist, dann sagen wir (K, w) ist ein Modell von φ, geschrieben $K, w \models \varphi$, mit

$K, w \models a$	genau dann, wenn	$a \in V(w)$
$K, w \models \neg\varphi$	genau dann, wenn	$K, w \not\models \varphi$
$K, w \models \varphi \wedge \psi$	genau dann, wenn	$K, w \models \varphi$ und $K, w \models \psi$
$K, w \models \varphi \vee \psi$	genau dann, wenn	$K, w \models \varphi$ oder $K, w \models \psi$
$K, w \models \varphi \Rightarrow \psi$	genau dann, wenn	$K, w \not\models \varphi$ oder $K, w \models \psi$
$K, w \models \Box\varphi$	genau dann, wenn	für alle $w' \in R(w)$ gilt $(K, w') \models \varphi$
$K, w \models \Diamond\varphi$	genau dann, wenn	es ein $w' \in R(w)$ gibt, sodass $(K, w') \models \varphi$

Für eine Formel $\Box\varphi$ gilt $K, w \models \Box\varphi$ insbesondere genau dann, wenn für jede von w aus erreichbare Welt gilt, dass dort φ gültig ist. Auf duale Art und Weise gilt $K, w \models \Diamond\varphi$ genau dann, wenn es mindestens eine w aus erreichbare Welt gibt, in der φ gültig ist. Man beachte, wenn φ eine Formel ist, die kein $\Box$ (oder $\Diamond$) enthält, dann ist (K, w) ein Modell von φ genau dann, wenn w ein Modell von φ im Sinne der Aussagenlogik ist. Die Besonderheit der Kripke-Semantik, respektive der Kripke-Strukturen, tritt also erst zutage, wenn die Formeln Modalitäten enthalten.

Für unsere Kripke-Struktur K_{pq} und die Welt w_1 können wir beobachten, dass dort $K_{\text{pq}}, w_1 \models \Box\Diamond Q$ gilt. Die Auswertung, um zu zeigen, dass $K_{\text{pq}}, w_1 \models$

$\Box\Diamond Q$ gilt, ist wie folgend:

$K_{\mathrm{pq}}, w_1 \models \Box\Diamond Q$ gilt, wenn $K_{\mathrm{pq}}, w_1 \models \Diamond Q$ und $K_{\mathrm{pq}}, w_2 \models \Diamond Q$

$K_{\mathrm{pq}}, w_1 \models \Diamond Q$ gilt, wenn $K_{\mathrm{pq}}, w_1 \models Q$ oder $K_{\mathrm{pq}}, w_2 \models Q$

$K_{\mathrm{pq}}, w_1 \models Q$ gilt

$K_{\mathrm{pq}}, w_2 \models \Diamond Q$ gilt, wenn $K_{\mathrm{pq}}, w_2 \models Q$, $K_{\mathrm{pq}}, w_3 \models Q$ oder $K_{\mathrm{pq}}, w_4 \models Q$

$K_{\mathrm{pq}}, w_2 \models Q$ gilt nicht

$K_{\mathrm{pq}}, w_3 \models Q$ gilt

Man beachte, dass die Formel $\Box\Diamond Q$ aber nicht in jeder Welt von K_{pq} gültig ist. Da von w_4 aus keine Welt erreichbar ist ($R(w_4) = \emptyset$), lässt sich einfach sehen, dass auch $K_{\mathrm{pq}}, w_4 \not\models \Diamond Q$ gilt. Konsequenterweise erhält man, dass auch $K_{\mathrm{pq}}, w_2 \not\models \Box\Diamond Q$ gelten muss.

Wir erweitern die Semantik auf alle Kripke-Strukturen wie folgt. Für jede Formel φ und jede Kripke-Struktur $K = \langle W, R, V\rangle$ definiert man:

$$K \models \varphi \quad \text{genau dann, wenn} \quad K, w \models \varphi \text{ für alle } w \in W$$

$$\models \varphi \quad \text{genau dann, wenn} \quad K \models \varphi \text{ für alle Kripke-Strukturen } K$$

Beispielsweise ist die Formel $\Box P \Rightarrow Q$ in jeder Welt von K_{pq} gültig. Somit gilt auch $K_{\mathrm{pq}} \models \Box P \Rightarrow Q$. Wie oben bereits gesehen haben gilt $K_{\mathrm{pq}}, w_2 \not\models \Box\Diamond Q$. Folglich muss auch $K_{\mathrm{pq}} \not\models \Box\Diamond Q$ und $\not\models \Box\Diamond Q$ gelten.

3.2.3 Anmerkungen zur Modallogik

Wir haben Modallogik aus der Perspektive der *Überzeugungen* betrachtet, also zur Modellierung von Informationen, die ein Agent *subjektiv* als wahr einstuft, aber nicht notwendigerweise *objektiv* wahr sein müssen. Dies wird auch als die *doxastische* Deutung bezeichnet. Darüber hinaus lassen sich $\Box$ und $\Diamond$ auch anders interpretieren. Manchmal wird explizit zwischen Wissen und Überzeugungen unterschieden. Bezieht man sich auf (objektives) Wissen, so spricht man von der *epistemischen* Deutung. In dieser steht $\Box\varphi$ für „Es ist sicher, dass φ gilt“, und $\Diamond\varphi$ steht für „φ könnte möglicherweise wahr sein“. In der *temporalen* Deutung, steht $\Box\varphi$ für „φ gilt zu jeder Zeit“, und $\Diamond\varphi$ steht für „φ gilt zu mindestens einem Zeitpunkt“. Die Deontik ist die Lehre von Pflichten, Normen und deren Bedeutung. Die deontische Interpretation von $\Box\varphi$ ist „Es ist geboten, das φ“, und $\Diamond\varphi$ steht für „Es ist erlaubt, dass φ“.

Für die verschiedenen Deutungen wird dieselbe oder eine ähnliche Semantik benutzt. Je nach Deutung und Anwendung werden aber andere Eigenschaften eingefordert. Forscher in der Modallogik ordnen einigen Eigenschaften festgelegte Namen zu, da diesen besonderen Bedeutungen zukommen:

(K) $\Box(\varphi \Rightarrow \psi) \Rightarrow (\Box\varphi \Rightarrow \Box\psi))$

Wenn der Agent überzeugt ist, dass ψ aus φ folgt, dann ist der Agent auch von ψ überzeugt, wenn er von φ überzeugt ist.

(D) $\Box\varphi \Rightarrow \Diamond\varphi$

Ist der Agent von φ überzeugt, so hält der Agent φ auch für möglich.

(T) $\varphi \Rightarrow \Diamond\varphi$

Wenn φ gilt, dann hält der Agent φ für möglich.

(4) $\Box\varphi \Rightarrow \Box\Box\varphi$

Wenn der Agent überzeugt ist, dass φ gilt, so ist der Agent davon überzeugt, dass er davon überzeugt ist, dass φ gilt.

(B) $\varphi \Rightarrow \Box\Diamond\varphi$

Wenn φ gilt, dann ist der Agent davon überzeugt, das er φ für möglich hält.

Die obige Eigenschaften sind universell zu verstehen. Das heißt, dass wenn man davon spricht, dass beispielsweise (K) gilt, dann meint man, dass für alle Formeln φ und ψ die Bedingung in (K) zutrifft. Häufig werden Eigenschaften gruppiert, so zum Beispiel die Mengen $\mathbf{S5} = \{(\mathrm{K}), (\mathrm{T}), (\mathrm{B}), (4)\}$ und $\mathbf{KD4} = \{(\mathrm{K}), (\mathrm{D}), (4)\}$. So gilt **S5** als gute Menge von Eigenschaften für die epistemische Deutung und **KD4** gilt als Grundlage für die doxastische Deutung. Man kann obige Eigenschaften garantieren, indem man die Kripke-Semantik aus Abschnitt 3.2.2 nutzt und sich zusätzlich auf bestimmte Kripke-Strukturen beschränkt. So ist beispielsweise die Eigenschaft (4) genau dann erfüllt, wenn man nur Kripke-Strukturen betrachtet, bei denen die Erreichbarkeitsrelation transitiv[6] ist. Weitere Informationen findet man in einer der exzellenten Einführungen in die Modallogik [27].

Wegen der Vielzahl an verschiedenen Deutungen und gleichzeitigen konzeptuellen Klarheit, finden Modallogiken in vielen Bereichen Anwendung, bzw. bilden die theoretische Grundlage. So hat die temporale Interpretation eine breite Anwendung in der Verifikation von Systemen. Die Formeln stehen dabei für

6 Transitivität: $(w, w') \in R$ und $(w', w'') \in R$ implizieren $(w, w'') \in R$

gewünschte Eigenschaften von Systemen und Systeme werden als (z.T. unendliche) Kripke-Strukturen modelliert. Kann man zeigen, dass die Kripke-Struktur ein Modell der Formel ist, so hat man implizit gezeigt, dass das System die gewünschte Eigenschaft hat.

3.3 Wissensdynamik

In diesem Abschnitt widmen wir uns dem Verändern der Wissensbasis durch neue Informationen, was wir auch als *Wissensänderung* bezeichnen. Als Wissensrepräsentationsformalismus benutzen wir dazu wieder die Aussagenlogik und unsere *Wissensbasis* K besteht daher aus einer Menge aussagenlogischer Formeln. Da wir bisher formal die semantischen Begriffe der Aussagenlogik nur auf einzelnen Formeln definiert haben, erweitern wir diese nun auf Wissensbasen, also Mengen von Formeln, wie folgt. Zunächst erinnern wir uns, dass eine Formel ψ aus ϕ folgt, geschrieben $\phi \models \psi$, wenn jedes Modell von ϕ ein Modell von ψ ist. Eine Interpretation ω ist dann ein Modell einer Wissensbasis K, falls ω ein Modell jeder Formel in K ist. Aus K folgt dann eine Formel ψ, wieder geschrieben $K \models \psi$, genau dann, wenn jedes Modell von K auch ein Modell von ψ ist. Eine Wissensbasis K wird als inkonsistent bezeichnet, falls es kein Modell für K gibt.

3.3.1 Beispiel

Wir beginnen mit einem einleitenden Beispiel von Peter Gärdenfors [12]:

> Eine Zoowärterin möchte gerne ein neues Tier für ihren Zoo kaufen. Nach sorgfältigem Überlegen kommt sie zu dem Entschluss einen neuen Schwan zu bestellen. Schlussendlich bestellt die Zoowärterin einen Schwan aus Schweden. Die Zoowärterin hat implizite Annahmen über den bestellten Schwan getroffen, die sich aus ihrem Hintergrundwissen ergeben. Aufgrund ihrer biologischen Vorbildung weiß die Zoowärterin, dass europäische Schwäne weiß sind. Auch weiß die Zoowärterin, dass Schweden ein Teil von Europa ist. Die Zoowärterin schließt daraus, dass der bestellte Schwan weiß sein wird. Als der neue Schwan eintrifft, stellt sich jedoch heraus, dass das Tier schwarz ist. Die Zoowärterin ist sich nun der Situation ausgesetzt einer ihrer Vorannahmen aufzugeben, da diese zusammen zu einer Inkonsistenz führen.

Betrachten wir das Beispiel nun etwas formaler mit den Mitteln der Aussagenlogik. Das Hintergrundwissen der Zoowärterin können wir als eine Menge von aussagenlogischen Formeln darstellen. Als Propositionen nutzen wir Schwan, $\mathsf{Schweden}$, Europa und $\mathsf{Weiß}$, wobei Schwan für die Aussage „Das Tier ist ein Schwan" steht; $\mathsf{Schweden}$ steht für die Aussage „Das Tier kommt aus Schweden"; Europa steht für die Aussage „Das Tier ist aus Europa"; und $\mathsf{Weiß}$ steht für die Aussage „Das Tier ist weiß". Die Zoowärterin weiß, dass es sich um einen Schwan aus $\mathsf{Schweden}$ handelt. Die Formel $\mathsf{Schweden} \Rightarrow \mathsf{Europa}$ drückt aus, dass alle schwedischen Schwäne auch europäische Schwäne sind. Die Formel $\mathsf{Schwan} \wedge \mathsf{Europa}) \Rightarrow \mathsf{Weiß}$ drückt aus, dass alle europäischen Schwäne weiß sind. Wir erhalten die folgende Wissensbasis, die das Wissen der Zoowärterin vor dem Eintreffen des Tieres modelliert:

$$K_{\mathsf{Zoo}} = \{\mathsf{Schwan}, \mathsf{Schweden}, \mathsf{Schweden} \Rightarrow \mathsf{Europa}, (\mathsf{Schwan} \wedge \mathsf{Europa}) \Rightarrow \mathsf{Weiß}\}$$

Mithilfe der Deduktion können wir $\mathsf{Weiß}$ aus K_{Zoo} schlussfolgern, d. h., es gilt $K_{\mathsf{Zoo}} \models \mathsf{Weiß}$.

3.3.2 Revision

Wir betrachten weiterhin die Wissensbasis K_{Zoo} und erhalten bei Erhalt des Schwans nun die neue Information, dass der Schwan schwarz ist, vereinfacht modelliert durch $\beta = \neg\mathsf{Weiß}$. Fügen wir diese neue Information einfach unserer Wissensbasis hinzu, aktualisieren unsere Wissensbasis also zu $K'_{\mathsf{Zoo}} = K_{\mathsf{Zoo}} \cup \{\beta\}$, so erhalten wir eine Inkonsistenz (K'_{Zoo} besitzt kein Modell). Damit ist K'_{Zoo} für uns nutzlos geworden, da wir aus einer inkonsistenten Wissensbasis alles schlussfolgern können: es gilt $K'_{\mathsf{Zoo}} \models \phi$ für alle Formeln ϕ.[7] So lässt sich aus der inkonsistenten Wissensbasis K'_{Zoo} ableiten, dass jeder schwedische Schwan europäisch ist ($K_{\mathsf{Zoo}} \cup \{\beta\} \models \mathsf{Schweden} \Rightarrow \mathsf{Europa}$), aber gleichzeitig auch ableiten, dass jeder schwedische Schwan nicht europäisch ist ($K_{\mathsf{Zoo}} \cup \{\beta\} \models \mathsf{Schweden} \Rightarrow \neg\mathsf{Europa}$).

Um in die Wissensbasis K_{Zoo} der Zoowärterin die neue Information β zu integrieren, müssen wir also die Inkonsistenz *auflösen*, d. h., Informationen aus $K_{\mathsf{Zoo}} \cup \{\beta\}$ aufgeben. Dabei sollten wir allerdings der neuen Information β Priorität einräumen, da dies neue Information darstellt von dessen Wahrheitsgehalt wir überzeugt sind. Wissensänderungsprozesse, die die Priorität auf die Integration

7 Dieses Phänomen heißt im Lateinischen *ex falso quodlibet* (*aus Falschem folgere alles*) und liegt darin begründet, dass $K \models \phi$ ja genau dann gilt, wenn jedes Modell von K auch Modell von ϕ ist; wenn es keine Modelle von K gibt, dann gilt dies trivial und unabhängig von ϕ für alle ϕ.

der neuen Information legen, werden als *Revision* bezeichnet. Formal beschreiben wir Wissensänderungen einer Wissensbasis K durch eine neue Information α als Funktion $\star$; $(K, \alpha) \mapsto K \star \alpha$, die die *Apriori* Wissensbasis und α auf den neuen Wissenszustand $K \star \alpha$ abbildet. Obige Bedingungen können wir als Prinzipien für Revisionen festhalten:

(Erfolg)	Es gilt $K \star \alpha \models \alpha$
(Konsistenz)	Wenn α konsistent ist, dann ist auch $K \star \alpha$ konsistent

Wenn wir an diesen Prinzipien festhalten, ergeben sich verschiedene Möglichkeiten, wie $K_{\mathsf{Zoo}} \star \beta$ aussehen könnte. Eine Änderung, bei der die Prinzipien (Erfolg) und (Konsistenz) erfüllt sind, ist, beispielsweise, die Änderung von K_{Zoo} zu

$$K_{\mathsf{Zoo}} \star_1 \beta = \{\mathsf{Schwan}, \mathsf{Schweden} \Rightarrow \mathsf{Europa}, (\mathsf{Schwan} \wedge \mathsf{Europa}) \Rightarrow \mathsf{Weiß}\} ,$$

in welchem die Zoowärterin die Überzeugung aufgibt, dass es sich um einen schwedischen Schwan handelt ($\mathsf{Schweden}$).

3.3.3 Das Prinzip der Relevanz

Jede Formel in K_{Zoo} trägt zum Konflikt in $K_{\mathsf{Zoo}} \cup \{\beta\}$ bei. Das heißt, wenn $\varphi \in K_{\mathsf{Zoo}}$ nicht in K_{Zoo} wäre, dann gäbe es auch keinen Konflikt. Aus einer rein logischen Perspektive ist es valide, jede Formel $\varphi \in K_{\mathsf{Zoo}}$ die zu dem Konflikt beiträgt, nicht in das Ergebnis der Revision $K_{\mathsf{Zoo}} \star \beta$ mit aufzunehmen. Wir haben aber bisher noch nicht besprochen, wie wir mit Formeln umgehen, die nicht zum Konflikt beitragen.

Erweitern wir das Beispiel. Nehmen wir an, dass die Zoowärterin einen männlichen schwedischen Schwan bestellt hat. Das heißt, die zusätzliche Annahme ist, dass es sich um einen männlichen Schwan handelt. Wir führen die Proposition Mann ein, die für „der Schwan ist männlich“ steht. Wir können die anfängliche Wissensbasis der Zoowärterin in diesem Szenario wie folgend modellieren:

$$\begin{aligned} K_{\mathsf{Zoo}}^{\mathsf{alt}} = \{&\mathsf{Mann},\ \mathsf{Schwan},\ \mathsf{Schweden}, \\ &\mathsf{Schweden} \Rightarrow \mathsf{Europa},\ (\mathsf{Schwan} \wedge \mathsf{Europa}) \Rightarrow \mathsf{Weiß}\} \end{aligned}$$

Revidiert man nun $K_{\mathsf{Zoo}}^{\mathsf{alt}}$ mit der Formel $\beta = \neg\mathsf{Weiß}$, so ist nicht plausibel aufzugeben, dass es sich um einen männlichen Schwan handelt (Mann). Daher sollte, egal wie die Revision durchgeführt wird, $\mathsf{Mann} \in K_{\mathsf{Zoo}}^{\mathsf{alt}} \star \beta$ gelten.

Allgemein fordert man, das Wissensänderungen sich so verhalten, dass sie nur relevantes Wissen aufgeben. Dies nennt man das *Prinzip der Relevanz*, das auf

Sven Ove Hansson zurückgeht [13], und das wir hier nur informell beschreiben:

(Prinzip der Relevanz) Wenn eine Formel bei einer Wissensänderung aufgeben wird, dann muss dies begründet sein durch einen Konflikt mit den Zielen der Wissensänderung.

In der Literatur zu Wissensänderungen findet man häufig auch das *Prinzip der minimalen Änderungen*, das fordert, dass Wissensänderungen stets minimale Änderungen am Ausgangswissen durchführen sollten. Dieses Prinzip geht historisch dem Prinzip der Relevanz voraus, in den meisten Kontexten ist das Prinzip der Relevanz aber allgemeiner und impliziert das Prinzip der minimalen Änderungen.

3.3.4 Algorithmische Konstruktion von Revisionsergebnissen

Wir betrachten nun, wie man konstruktiv ein Revisionsergebnis erhalten kann. Dabei wenden wir eine Konstruktion an, die auf der *Maxichoice-Kontraktion* von Alchourron und Makinson [1] basiert. Die Eingabe dieses Verfahrens besteht aus einer Wissensbasis K und einer Formel α. Das Verfahren gibt eine Revision von K durch α aus, bei der die Prinzipien (Erfolg), (Konsistenz) und das (Prinzip der Relevanz) erfüllt sind. Zentraler Aspekt des Verfahrens ist es eine Teilmenge M der Wissensbasis K zu bestimmen, für die gilt:

– $\neg\alpha$ ist keine Schlussfolgerung aus M, das heißt, $M \not\models \neg\alpha$;
– M ist maximal mit diesen Eigenschaften, das heißt, für alle M' mit $M \subsetneq M' \subseteq K$ gilt $M' \models \neg\alpha$.

Das Revisionseregebnis ist dann $M \cup \{\alpha\}$. Ein naiver Algorithmus zählt (absteigend) alle Teilmengen X von K nach Größe sortiert auf, beginnend bei der größten Teilmenge (die K selbst ist). In jeden Schritt testen wir ob $X \not\models \neg\alpha$ gilt. Falls nein, dann machen wir mit der nächsten Teilmenge von K weiter. Falls ja, setzen wir $M := X$ und stoppen das Aufzählen. Als Ergebnis geben wir dann $M \cup \{\alpha\}$ aus.

Wenden wir als Beispiel das Verfahren auf die Wissensbasis $K^{\mathsf{alt}}_{\mathsf{Zoo}}$ an, um eine Revision $K^{\mathsf{alt}}_{\mathsf{Zoo}} \star \beta$ zu berechnen. Wir setzen $X' := K^{\mathsf{alt}}_{\mathsf{Zoo}}$ als ersten Schritt. Es gilt $K^{\mathsf{alt}}_{\mathsf{Zoo}} \models \neg\beta$ und wir fahren fort. Im nächsten Schritt zählen wir $X' := K^{\mathsf{alt}}_{\mathsf{Zoo}} \setminus \{\mathsf{Mann}\}$ auf. Auch für dieses X' gilt $X' \models \neg\beta$. Nun zählen wir $X' := K^{\mathsf{alt}}_{\mathsf{Zoo}} \setminus \{\mathsf{Schwan}\}$ auf und stellen fest, dass $X' \not\models \neg\beta$ gilt. Das Ergebnis der Revision nach obigen Verfahren ist damit:

$$\begin{aligned}&(K^{\mathsf{alt}}_{\mathsf{Zoo}} \setminus \{\mathsf{Schwan}\}) \cup \{\neg\mathsf{Weiß}\}\\ &\quad = \{\mathsf{Mann}, \mathsf{Schweden}, \mathsf{Schweden} \Rightarrow \mathsf{Europa}, (\mathsf{Schwan} \wedge \mathsf{Europa}) \Rightarrow \mathsf{Weiß}, \neg\mathsf{Weiß}\}\end{aligned}$$

Es gibt viele weitere (insbesondere effektivere) Methoden eine Revision zu konstruieren. Die klassischen Ansätze folgen obigen Ansatz und nutzen maximale Teilmenge der Wissensbasis K wie oben beschrieben. Weitere Konstruktionsmethoden ordnen die Formeln anhand ihrer *epistemischen Verwurzelung*. Das Ergebnis der Revision sind dann diejenigen Formeln, die am stärksten „verwurzelt" und mit α kompatibel sind. Semantische Konstruktionsmethoden bilden Ordnungen über den Interpretationen, die dann als mögliche Welten (Englisch „*possible worlds*") verstanden werden. Das Revisionsergebnis von K mit α entspricht semantisch den am meisten präferierten möglichen Welten, die auch Modelle von α sind. Die Forschung enthält eine Vielzahl von Resultaten, die genau die Eigenschaften dieser verschiedenen Konstruktionsansätze charakterisieren.

3.3.5 Verschiedene Operationen der Wissensänderungen

Es gibt verschiedene Arten der Wissensänderungen und die Formalisierung dieser folgt dem Schema der Revision. Ausgangslage ist eine Wissensbasis K und eine neue Information α, die die Wissensänderungsoperation in eine neue Wissensbasis $K \circ \alpha$ überführen. Hier geben wir einen kleinen Überblick über wichtige Wissensänderungsarten:

– Die **Revision** [2] haben wir weiter oben ausführlich behandelt. Es handelt sich um eine Wissensänderung, bei der α in das bestehende Wissen K so integriert wird, dass α im Endzustand geglaubt wird, und Inkonsistenzen (wenn möglich) aufgelöst werden.
– Bei der **Kontraktion** handelt es sich um eine Wissensänderung, bei der die Information α aufgegeben werden soll, und zwar so, dass keine neuen Informationen hinzugefügt werden [2].
– Bei der **Credibility-Limited Revision** (Deutsch: „Glaubwürdigkeitsbeschränkte Revision") handelt es sich um eine Art der Wissensänderung, die eine Revision nur durchführt, wenn die neue Information α als glaubwürdig eingestuft wird [15].
– Bei der **Shielded Contraction** (Deutsch: „Abgeschirmte Kontraktion") handelt es sich um das Analog der Credibility-Limited Revision für die Kontraktion [9]. Hier wird eine Kontraktion von α nur durchgeführt, wenn der Agent es als valide betrachtet, die Information α aufzugeben.
– Wie bei der Revision, handelt es sich bei **Update** [16] um eine Art der Wissensänderung, bei der α in das bestehende Wissen K so integriert wird, das α im Endzustand geglaubt wird und Inkonsistenzen (wenn möglich) aufgelöst werden. Der Unterschied besteht darin, dass man für eine Revision annimmt,

dass sich die Welt nicht geändert hat. Die neue Information α beschreibt also weitere genauere Informationen, während K valide, aber potenziell ungenauere, Informationen enthält. Beim Update nimmt man jedoch an, dass die Welt sich geändert haben könnte. Dies hat drastische semantische Unterschiede, insbesondere, da die Informationen in der Wissensbasis K potenziell nicht mehr aktuell sind.

– Beim **Improvement** (Deutsch: „Verbessern“) handelt es sich um eine Verallgemeinerung der Revision, bei der die neue Information α nur als Anlass genommen wird, die Glaubwürdigkeit der Information α zu verbessern [17]. Konsequenterweise führt ein **Improvement** von α, nicht direkt dazu, dass α geglaubt wird.
– Beim **Decrement** (Deutsch: „Dekrementierung“) handelt es sich, im Prinzip, um die duale Operation zum **Improvement** [24]. Ein Dekrementierung von α führt dazu, dass die Glaubwürdigkeit der Information α verringert wird. Konsequenterweise führt ein **Decrement** von α, nicht direkt dazu, dass α aufgegeben wird.

Neben obigen Operatoren gibt es noch viele weitere Arten von Wissensänderungen, auf die wir hier nicht weiter eingehen, aber gerne auf Übersichtsarbeiten verweisen [13, 10].

3.3.6 Weitere Aspekte und Entwicklung

Wir erwähnen einige weitere Teilgebiete und Einsichten aus der Forschung zu Wissensänderungen.

Eine zentrale Einsicht des Forschungsfelds ist, dass sich mit rein logischen Betrachtungen nicht immer optimale Wissensänderungen durchführen lassen, sondern *extra-logische Informationen* nötig sind [11]. Betrachten wir noch einmal das Eingangsbeispiel. Wir können feststellen, dass für jede echte Teilmenge $K'_{\mathsf{Zoo}} \subsetneq K_{\mathsf{Zoo}}$ gilt, dass $K'_{\mathsf{Zoo}} \cup \{\beta\}$ aus einer rein logischen Perspektive ein valides Ergebnis für $K_{\mathsf{Zoo}} \star \beta$ ist (es sind (Erfolg), (Konsistenz) und das (Prinzip der Relevanz) erfüllt). Eine dieser Lösungen ist jedoch

$$K_{\mathsf{Zoo}} \star_2 \beta = \{\mathsf{Schwan}, \mathsf{Schweden}, (\mathsf{Schwan} \wedge \mathsf{Europa}) \Rightarrow \mathsf{Weiß}\} \ ,$$

in der die Zoowärterin die Überzeugung aufgibt, dass schwedische Schwäne stets auch europäische Schwäne sind ($\mathsf{Schweden} \Rightarrow \mathsf{Europa}$). Aus der Perspektive des gesunden Menschenverstands macht es aber keinen Sinn, eine generische Schlussfolgerung wie $\mathsf{Schweden} \Rightarrow \mathsf{Europa}$ aufzugeben. Insbesondere weil man alternativ auch nur das spezifischere Wissen aufgeben kann, dass es sich um einen schwedischen Schwan handelt (Schwan). Die Information, welche Formeln aus K „wertvoller“ sind als andere Formeln aus K, kann aber nicht die Logik liefern, sondern muss extern bereitgestellt werden.

Ein sehr aktives Forschungsfeld ist das Gebiet der *Iterierten Wissensänderung*. In diesem Gebiet beschäftigt man sich mit der Frage von Konstruktionen und Eigenschaften von mehreren Änderungen nacheinander, also Folgen von Änderungen der Art $K \star \alpha \star \beta$ und so weiter. Besonderer Bedeutung in diesem Forschungsgebiet kommt Überzeugungen über *Konditionalen* zu, wobei ein Konditional häufig als Information der Form $(\beta \mid \alpha)$ formalisiert wird und beispielsweise dafür stehen kann das β *normalerweise* unter der Prämisse α gilt. Über dem Ramsey-Test [26], sind Konditionale direkt mit Wissensänderungen verbunden:

$$(\beta \mid \alpha) \text{ wird akzeptiert von } K\text{, genau dann wenn } \beta \in K \star \alpha \text{ gilt.}$$

Iterierte Wissensänderungen beeinflussen damit auch die Akzeptanz von Konditionalen, da in $K \star \alpha$ eventuell andere Konditionale akzeptiert werden als in K. Eine der wichtigsten Einsichten des Forschungsfeldes ist, dass man den Effekt iterativer Änderungen auf die Akzeptanz von Konditionalen einschränken sollte, aber im Gegensatz zum Prinzip der minimalen Änderungen, nicht minimieren sollte [8].

Ein weiteres Forschungsthema, ist der Bereich der *multiplen Wissensänderungen* [22, 21]. Hier führt man eine Wissensänderung nicht nur mit einer Information α, sondern mit Menge von Informationen $A = \{\alpha_1, \ldots, \alpha_n\}$ durch. Maßgeblich werden zwei Semantiken für solche Arten von Änderungen betrachtet. Die erste Semantik ist die *choice* Semantik (Deutsch: „Auswahlsemantik“), bei der man verlangt, dass die Wissensänderung für eine der Formeln in A erfolgreich ist. Die zweite Semantik ist die *package* Semantik (Deutsch: „Paketsemantik“), bei der man verlangt, dass die Wissensänderung für alle Formeln in A erfolgreich ist. Eine Revision von K mit der Menge A ist unter der Auswahlsemantik erfolgreich, wenn für eine Formel $\alpha_i \in A$ gilt, dass $\alpha_i \in K \star A$. Für die Paketsemantik ist die Revision von K mit A erfolgreich, wenn für alle Formel $\alpha_i \in A$ gilt, dass $\alpha_i \in K \star A$.

Zuletzt wollen wir erwähnen, dass es Bestrebungen gibt, die verschiedenen Arten der Wissensänderungen aus einer vereinheitlichenden Perspektive zu betrachtet. Wesentlich ist hierbei der Ansatz der *Descriptor Revision* nach Hansson [14] zu nennen. Die Grundidee dieses Ansatzes ist es, dass oben skizzierte Schema eines Wissensänderungsoperator so zu erweitern, dass statt nur die Information α bei der Änderung miteinzubeziehen, wir einen sogenannten *Deskriptor* übergeben. Ein Deskriptor Ψ spezifiziert dabei eine Erfolgsbedingung für den Wissensänderungsprozess durch die Angabe von Glaubensprädikaten der Form $\mathbf{B}\varphi$. Beispielsweise drückt der Deskriptor $\neg\mathbf{B}\beta \wedge (\mathbf{B}\alpha \vee \mathbf{B}\neg\alpha)$ aus, dass die

Wissensänderung als erfolgreich zählen soll, wenn nicht an β geglaubt wird, und an α oder $\neg\alpha$ geglaubt wird. Prinzipiell können wir mit einem Deskriptor $\mathbf{B}\alpha$ ausdrücken, dass eine Revision mit α durchgeführt werden soll; und mit $\neg\mathbf{B}\alpha$ wird ausdrückt, dass eine Kontraktion von α durchgeführt werden soll. Dies lässt sich auf andere Wissensänderungsoperationen erweitern.

4 Ausblick

Wissensrepräsentation und -verarbeitung ist seit seiner Begründung ein zentrales Forschungsfeld der Künstlichen Intelligenz und hat sich in verschiedene Teilgebiete ausdifferenziert. Die großen Organisationen der Informatik, wie die *Association for Computing Machinery* (ACM, Deutsch: „Amerikanische Gesellschaft für Informatik") und auch die großen Konferenzen der Künstlichen Intelligenz, wobei insbesondere die *International Joint Conference on Artificial Intelligence* (IJCAI, Deutsch: „Internationale Konferenz für Künstliche Intelligenz") zu nennen ist, haben Klassifikationen der Teilgebiete eingeführt. Weiter oben haben wir bereits die Bereiche der **Beschreibungslogik**, der **Wissenslogiken** und die **Wissensdynamik** behandelt. Wir geben einen kurzen Ausblick über die weiteren Teilgebiete der Wissensrepräsentation und -verarbeitung:

Automatische Inferenz. Dieser Forschungszweig beschäftigt sich mit Verfahren und Prinzipien des automatisierten Schließens. Wichtige Themen sind das Finden und Verifizieren von Beweisen. *(siehe auch Kapitel 7)*

Semantische Netze. In diesen Bereich beschäftigt man sich mit Formalismen und Methoden um semantische Informationen zu sammeln. Als Standard haben sich dabei Wissensgraphen etabliert, für die es zahlreiche Erweiterungen gibt. *(siehe auch Kapitel 21)*

Ontologien. Forscher beschäftigen sich mit den Voraussetzungen und Schwierigkeiten um Wissensdatenbanken aufzubauen, die insbesondere auch strukturelle Informationen enthalten. Besonders zu nennen ist hier die Web Ontology Language (OWL), die auch als Standardsprache für semantische Netze fungiert. *(siehe auch [3])*

Fallbasiertes Schließen. Diese spezielle Methodologie des Schließen zeichnet sich dadurch aus, dass es maßgeblich *Erfahrungen-basierendes* Schließen realisiert. Das heißt, wenn aus den vorliegenden Daten (dem *Fall*) ein Schluss gezogen werden soll, werden Fälle aus der Vergangenheit mit dem aktuellen Fall verglichen. Basierend auf der Nähe zu den anderen Fällen, wird ein

Entschluss getroffen. Der neue Fall wird in die bestehende Fallbasis integriert. *(siehe auch Kapitel 11)*

Nichtmonotones Schließen. Diese Art des Schließens ist begründet in der Beobachtung, dass in der Praxis bereits gezogene Schlüsse zurückgenommen werden können. Logiken, die bei Hinzunahme neuer Informationen eine Rücknahme vorheriger Schlussfolgerungen erlauben, nennt man *nicht-monoton*. Klassische Logiken wie die Aussagenlogik sind monoton, sodass Forscher in diesem Feld Verfahren und Logiken untersuchen, die nicht-monoton sind. *(siehe auch Kapitel 8)*

Argumentation. Dieses Teilgebiet beschäftigt sich mit der Formalisierung und dem automatischen Schlussfolgern mit Argumenten und Gegenargumenten. Dies beinhaltet die Repräsentation von Argumenten und Feststellung von *Akzeptanz* von Mengen von Argumenten. *(siehe auch [4])*

Unsicheres, unpräzises und unscharfes Wissen. Oft liegt kein gesichertes Faktenwissen vor, aber Agenten sind trotzdem der Situation ausgesetzt unter diesen Bedingungen handeln und schließen zu müssen. Daher beschäftigt sich dieses Feld grundlegend mit allen Arten von Unsicherheiten und wie damit umzugehen ist. *(siehe auch Kapitel 9)*

Kausalität. Dieses Feld beschäftigt sich mit der Repräsentation und Aspekten von Ursache-Wirkungs-Beziehungen, Kausalitäten genannt. Eine Besonderheit spielen dabei *kontrafaktische Konditionale*, also Sätze der Art „Wenn es regnen würde, dann hätte ich einen Regenschirm mitgenommen". Unter bestimmten Voraussetzungen beschreiben kontrafaktische Konditionale diese Kausalitäten. *(siehe auch [20])*

Präferenzbasiertes Schließen. Wissen kann auch Präferenzen unterliegt. So kann ein Agent manches Wissen für plausibler halten als anderes Wissen (beispielsweise kann man die Tatsache, dass man seinen Nachbarn im Supermarkt getroffen hat, als plausibler halten als die Tatsache, dass man eine berühmte Person im Supermarkt getroffen hat). Dieses Forschungsfeld untersucht Voraussetzungen, Semantiken und Methoden für Schlussfolgerungen, die Präferenzen miteinbeziehen. *(siehe auch [7])*

Logische Programmierung. Bei der logischen Programmierung handelt sich um ein Programmierparadigma, bei dem Ziele und Regeln eines Programms deklarativ beschrieben werden und eine Lösung automatisiert bestimmt werden kann. Typische Vertreter dieses Programmierparadigmas sind Prolog und die Antwortmengenprogrammierung. *(siehe auch [19])*

Schließen über Raum und Zeit. Agenten müssen beim Handeln und Schließen auch Zeit und Raum mit einbeziehen. Dieses Forschungsfeld beschäftigt sich mit grundlegenden Methoden, wie Raum und Zeit als Wissen repräsentiert werden können, aber auch damit, wie Wissen über Zeit und Raum das Schließen und Handeln beeinflusst. *(siehe auch Kapitel 10)*

Die Klassifikation der Teilgebiete der Wissensrepräsentation und -verarbeitung ist allerdings auch historischen Wandlungen unterlegen. Beispielsweise war das Gebiet des *Automatischen Planens* (siehe Kapitel 3) lange Zeit auch ein Teilgebiet der Wissensrepräsentation und -verarbeitung. Die Forschung und Anwendung von Methoden des automatischen Planens hat aber an so immenser Bedeutung gewonnen, dass dieses heutzutage als eigenes Forschungsgebiet behandelt wird.

Weitere Kapitel dieses Buches behandeln viele der oben genannten Themen mit mehr Details. So wird im nächsten Kapitel 7 automatische Inferenz tiefgehender diskutiert. Ebenso werden Nichtmonotones Schließen (Kapitel 8), Unsicheres, unpräzises und unscharfes Wissen (Kapitel 9), Wissen über Raum und Zeit (Kapitel 10), Fallbasiertes Schließen (Kapitel 11), sowie semantische Netze (Kapitel 21) explizit in diesem Buch behandelt.

5 Schlussbetrachtung

Dieses Kapitel gab einen kurzen Überblick über den Forschungsbereich der *Wissensrepräsentation und -verarbeitung.* Wir haben uns mit grundlegenden Eigenschaften der formalen Logik und insbesondere der Aussagenlogik beschäftigt und sind etwas tiefer in die Bereiche der Beschreibungslogiken, der Wissenslogiken und der Wissensdynamik eingestiegen. Schließlich haben wir im letzten Abschnitt weitere Teilbereiche dieses Forschungsbereichs angerissen.

Wie bereits der letzte Abschnitt angedeutet hat, haben wir in diesem Kapitel nur einen kleinen Ausschnitt aus dem Forschungsgebiet der *Wissensrepräsentation und -verarbeitung* betrachtet. Einen breiteren Einstieg in das Gebiet bieten beispielsweise die Bücher von Russel und Norvig [23], Brachman und Levesque [6], sowie von Beierle und Kern-Isberner [5]. Aktuelle Forschungsergebnisse werden auf der jährlich stattfindenden Konferenz KR (*International Conference on Knowledge Representation and Reasoning*), den allgemeinen internationalen KI-Konferenzen IJCAI (*International Joint Conference on Artificial Intelligence*), AAAI (die Konferenz der *Association for the Advancement of Artificial Intelligence*), ECAI (*European Conference on Artificial Intelligence*), sowie den Zeitschriften *Artificial Intelligence* und *Journal of Artificial Intelligence Research* vorgestellt.

Eines der spannendsten aktuellen Forschungsfragen behandelt die Integration von Methoden der *Wissensrepräsentation und -verarbeitung* mit Methoden des

Maschinellen Lernens und insbesondere des *Deep Learnings*. Während Methoden des zuletzt genannten Bereichs insbesondere in jüngster Zeit große Erfolge in einem breiten Spektrum verschiedenster Anwendungsfälle verzeichnen konnten, so arbeiten sie dennoch nach einem *Black-Box*-Prinzip und Klassifikationsergebnisse oder natürlichsprachliche Aussagen können kaum bis gar nicht *erklärt* oder belegt werden. Große Sprachmodelle, wie solche, die in ChatGPT verwendet werden, leiden beispielsweise unter *Halluzinationen*, also die Erzeugung von Aussagen, die in keinster Weise aus den Daten ableitbar sind. Dies liegt darin begründet, dass solche Modelle keine explizite Wissensrepräsentation besitzen und am Ende des Tages nur auf Häufigkeiten von zusammen vorkommenden Worten und Begriffen arbeiten. Gerade in sicherheitskritischen Bereichen, wie beispielsweise in der Medizin, müssen Entscheidungen aber klar nachvollziehbar und erklärbar sein. Methoden der *Wissensrepräsentation und -verarbeitung* basieren auf der Grundlage der formalen Logik und sind damit inhärent erklärbar und arbeiten faktenbasiert. Logische Schlussfolgerungen können explizit dargestellt und visualisiert werden und können damit der Nachvollziehbarkeit von Vorhersagen dienen und das Vertrauen in das System stärken. Eine Kombination der Vorteile dieser beiden Forschungsbereiche, also die Möglichkeit der *skalierbaren datengetriebenen Modellbildung* des *Maschinellen Lernens* und der *erklärbaren faktengestützen Schlussfolgerungsmechanismen* der *Wissensrepräsentation und -verarbeitung*, stellt damit eine vielversprechende Richtung für die zukünftige Erforschung von künstlicher Intelligenz dar.

Danksagung: Unser besonderer Dank gilt den Autoren der Vorauflage dieses Kapitels, Bernhard Nebel und Stefan Wölfl.

Literaturverzeichnis

[1] Carlos E. Alchourron and David Makinsion. On the logic of theory change: Contraction functions and their associated revision functions. *Theoria*, 48(1):14–37, 1982.

[2] Carlos E. Alchourrón, Peter Gärdenfors, and David Makinson. On the logic of theory change: Partial meet contraction and revision functions. *Journal of Symbolic Logic*, 50(2):510–530, 1985.

[3] Franz Baader, Diego Calvanese, Deborah L. McGuinness, Daniele Nardi, and Peter F. Patel-Schneider, editors. *The Description Logic Handbook: Theory, implementation, and applications.* Cambridge University Press, 2010.

[4] Pietro Baroni, Dov Gabbay, Massimilino Giacomin, and Leendert van der Torre, editors. *Handbook of Formal Argumentation*, volume 1. College Publications, 2018.

[5] Christoph Beierle and Gabriele Kern-Isberner. *Methoden wissensbasierter Systeme - Grundlagen, Algorithmen, Anwendungen.* Computational intelligence. SpringerVieweg, 6 edition, 2019.

[6] Ronald J. Brachman and Hector J. Levesque. *Knowledge Representation and Reasoning.* Elsevier, 2004.

[7] Felix Brandt, Vincent Conitzer, Ulle Endriss, Jérôme Lang, and Ariel D. Procaccia, editors. *Handbook of Computational Social Choice.* Cambridge University Press, 2016.

[8] A. Darwiche and J. Pearl. On the logic of iterated belief revision. *Artificial Intelligence*, 89:1–29, 1997.

[9] Eduardo L. Fermé and Sven Ove Hansson. *Shielded Contraction*, pages 85–107. Springer, Dordrecht, 2001.

[10] Eduardo L. Fermé and Sven Ove Hansson. *Belief Change - Introduction and Overview.* Springer Briefs in Intelligent Systems. Springer, 2018.

[11] Peter Gärdenfors. *Knowledge in flux : modeling the dynamics of epistemic states.* MIT Press Cambridge, 1988.

[12] Peter Gärdenfors. *Belief revision: An introduction*, page 1–28. Cambridge Tracts in Theoretical Computer Science. Cambridge University Press, 1992.

[13] Sven Ove Hansson. *A Textbook of Belief Dynamics: Theory Change and Database Updating.* Springer, 1999.

[14] Sven-Ove Hansson. *Descriptor Revision.* Springer, 2017.

[15] Sven Ove Hansson, Eduardo L. Fermé, John Cantwell, and Marcelo A. Falappa. Credibility limited revision. *Journal of Symbolic Logic*, 66(4):1581–1596, 2001.

[16] Hirofumi Katsuno and Alberto O. Mendelzon. On the difference between

https://doi.org/10.1515/9783110218091-002

updating a knowledge base and revising it. In James F. Allen, Richard Fikes, and Erik Sandewall, editors, *Proceedings of the 2nd International Conference on Principles of Knowledge Representation and Reasoning (KR 1991), Cambridge, MA, USA*, pages 387–394, 1991.

[17] Sébastien Konieczny and Ramón Pino Pérez. Improvement operators. In Gerhard Brewka and Jérôme Lang, editors, *Proceedings of the 11th International Conference on Principles of Knowledge Representation and Reasoning (KR 2008), Sydney, Australia*, pages 177–187. AAAI Press, 2008.

[18] Saul A. Kripke. Semantical analysis of modal logic i normal modal propositional calculi. *Mathematical Logic Quarterly*, 9(5-6):67–96, 1963.

[19] Vladimir Lifschitz. *Answer set programming.* Springer, 2019.

[20] Judea Pearl. *Causality: Models, Reasoning and Inference.* Cambridge University Press, 2nd edition, 2009.

[21] Pavlos Peppas. The limit assumption and multiple revision. *Journal of Logic and Computation*, 14(3):355–371, 2004.

[22] Hans Rott. Modellings for belief change: Prioritization and entrenchment. *Theoria*, 58(1):21–57, 1992.

[23] Stuart Russell and Peter Norvig. *Artificial Intelligence: A Modern Approach.* Pearson, 4 edition, 2020.

[24] Kai Sauerwald and Christoph Beierle. Decrement operators in belief change. In Gabriele Kern-Isberner and Zoran Ognjanovic, editors, *Proceedings of the 15th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU 2019), Belgrade, Serbia*, volume 11726 of *Lecture Notes in Computer Science*, pages 251–262. Springer, 2019.

[25] Manfred Schmidt-Schauß and Gert Smolka. Attributive concept descriptions with complements. *Artificial Intelligence*, 48(1):1–26, 1991.

[26] Robert Stalnaker. A theory of conditionals. In Nicholas Rescher, editor, *Studies in Logical Theory*, pages 98–112. Blackwell, 1968.

[27] Johan van Benthem. *Modal Logic for Open Minds.* Number 199 in CSLI Lecture Notes. CSLI Publications, 2010.